Black Box Correlations:

Locality, Noncontextuality, and Convex Polytopes

Master's Thesis

Marco Túlio Quintino

Supervisor: Marcelo Terra Cunha
Co-supervisor: Daniel Cavalcanti

Examination board: Daniel Joanathan
Marcelo França Santos

Master's thesis presented to *Programa de Pós-Graduação em Física da Universidade Federal de Minas Gerais* as a partial request for the title of Master in physics.

Semptember 2012
Belo Horizonte – Brazil

We are all in the gutter, but some of us are looking at the stars...

Oscar Wilde

Agradecimentos

Começo os agradecimentos pelo grande mestre Marcelo Terra Cunha, que desde as aulas de cálculo 1, consegue me esclarecer e estimular conceitos matemáticos, físicos e filosóficos. E mesmo fornecendo a liberdade para estudar o que eu quiser[1], soube os momentos certos de exigir foco.

A minha inspiradora Flávia, que além de cuidar de mim desde antes do mestrado, leu meus textos, ouviu a prévia de apresentações e me sugeriu escolhas sabias. Sua participação nos momentos felizes e difíceis é fundamental para mim. (L)

Ao grande amigo e colaborador Mateus, que desde a graduação, proporciona discussões estimulantes, essenciais para a existência desta dissertação.

Ao também amigo e colaborador (big) Marcelo, que sempre consegue acrescentar nas discussões com detalhes que passavam despercebido.

Ao Daniel, um dos responsáveis pelo meu interesse em não-localidade, agradeço pelas conversas empolgantes e esclarecedoras, por confiar em mim desde o primeiro contato e pelas grandes ajudas relacionadas ao doutorado.

A todo grupo do EnLight, em especial a Bárbara e a Gláucia, meninas responsáveis pelo o crescimento da não-localidade na UFMG, que também fizeram comentários importantes sobre o texto.

Ao pessoal do cafofo, Reinaldo, Tanus, Tche, Iemini[2], Debarba, pelas várias discussões de informação quântica, pela parceragem, pelas cervejas e por não me desejarem a morte[3].

Ao Rafael, fundamental para meus primeiros passos como pesquisador, agradeço a hospitalidade em Singapura, as referências importantes nas horas certas e as suas influências não-locais.

Ao Fernando (Brandão), pelas importantes conversas sobre música e mecânica quântica, e pelo apoio durante a escolha de um bom doutorado.

Ao Adán Cabello, pela confiança, empolgação e ideias que me motivam a crescer como pesquisador.

Aos vários amigos de viagens e congressos, os quais ainda que com pequenos comentários, acrescentaram bastante em minha formação.

Aos amigos feitos na física; Yulia, pelos importantes cafés, pelos treinos de parkour e por me levar para comer nas horas certas. Véronic, por conseguir decriptar meus textos[4], tornando-os legíveis e pelas várias boas conversas.

[1] Chegando a permitir/incentivar um estudo sério sobre a hipótese de Riemann.

[2] Iemino desu.

[3] Apesar de eu não conseguir garantir a validade desta afirmação para todo tempo $t \in \mathbb{R}$.

[4] Monografia, dissertação, etc.

Cobrex, pelo curso em semicondutores. Os obreiros Custelis e Mario Mazzoni. E ao hermano Ariel, que com poucos encontros tornou-se um grande amigo.

Aos amigos de Div; Rockmile, que desde tempos primórdios me surpreende com sua empolgação, carinho e pureza. Gigi, pelas conversas importantes nos dias importantes. Mah Fe, por ter me incentivado a estudar física/matemática e pelo carinho durante todo este tempo. Bode, que me acompanhou no começo de tudo, pelos ensinamentos e histórias. E meus três eternos amigos, Hermes, Pedro e Prunus[5], presentes em vários momentos importantes da vida.

Ao Gustavo, pelas sábias palavras e por me ajudar a compreender quem eu sou.

Ao pessoal daqui de casa[6], pelas musicas, casos e bebidas. Ao Leandrim e ao Brenin[7], por serem amigos confiáveis e presentes desde a época do GAAL. Ao rei PC, brother desde sua existência, companheiro de lutas, musicas, pilotagens e video games, eu agradeço pela confiança e verdadeira amizade durante todo este tempo[8]. Você vai fazer falta em Genebra. . .

Ao meu primo Dogruts[9], pelos conhecimentos de musica.

Ao meu pai, minha mãe e minha mainha, pelo apoio indispensável, e pelas várias vezes que estavam lá quando eu precisei. Sem eles este texto provavelmente nem teria começado.

Agradeço também ao programa de pós-graduação da física[10], ao CAPES e o CNPq pelo suporte financeiro e por proporcionar viagens.

[5]É mermão. . . é o Prunus!

[6]PC, Leandrim, Brenin e Frangão.

[7]Viadin.

[8]Agradeço também pelas comidas e biscoitos 0800.

[9]E além de primo, também arqui-inimigo durante as partidas de Mario Kart.

[10]Em especial aos atuais coordenadores Monken e Marcelo França.

Abstract

This dissertation studies nonlocal/contextual correlations and its generalisation from a *device independent* approach, where statistical data are themselves the object of study. For this, we explore a *black box* framework, in which a list of input buttons can be pressed to provide, with some probabilities, a list of outputs. In this formalism, we discuss general examples, analyse correlations that arise from quantum mechanics, and present all the inequalities characterise the noncontextual polytope for the n-cycle scenario. Moreover, we prove some efficiency requirements for nonlocality in an imperfect scenario, and use these results to propose a physical loophole free Bell test in an optical quantum system where photodetection and homodyne measurements are performed. Our findings contributes to the comprehension of nonlocal/contextual correlations and add to efforts towards feasible proposals for loophole-free Bell tests.

Resumo

Esta dissertação trabalha com correlações não-locais/contextuais e suas possíveis generalizações em uma abordagem *device independent*, onde apenas os resultados estatísticos são analisados. Para isso, exploramos uma "caixa preta", na qual diferentes botões de entrada podem ser apertados para que se obtenha, com dadas probabilidades, diferentes saídas. Neste formalismo, discutimos exemplos gerais, analisamos correlações oriundas da mecânica quântica, e apresentamos todas as desigualdades que caracterizam os politopos não-contextual do cenário n-ciclo. Nós também demonstramos algumas cotas de eficiência necessárias para se obter não-localidade em um cenário com imperfeições, depois usamos estes resultados para propor um teste de Bell livre de *loopholes* em um sistema óptico onde fotodeteção e medições homodinas são realizadas. Nossos resultados contribuem na compreensão de correlações não-locais/contextuais e na busca de testes de Bell livre de *loopholes*.

Contents

List of publications

Parts of this thesis is based on material published in the following papers:

- [1] **Maximal violations and efficiency requirements for Bell tests with photodetection and homodyne measurements**
  Marco Túlio Quintino, Mateus Araújo, Daniel Cavalcanti, Marcelo França Santos, Marcelo Terra Cunha
  Journal of Physics A Mathematical General 45, 21 (2012) 215308
  Section 4.4

- [2] **Tests of Bell inequality with arbitrarily low photodetection efficiency and homodyne measurements**
  Mateus Araújo, Marco Túlio Quintino, Daniel Cavalcanti, Marcelo França Santos, Adán Cabello, Marcelo Terra Cunha
  Phys. Rev. A 86, 3 (2012) 030101
  Section 4.4

- [3] **Realistic loophole-free Bell test with atom-photon entanglement**
  Colin Teo, Mateus Araújo, Marco Túlio Quintino, Jiří Minář, Daniel Cavalcanti, Valerio Scarani, Marcelo Terra Cunha, Marcelo França Santos
  submitted – Nature Communications 4, (2013)
  Section 4.4.10

- [4] **Complete characterization of the n-cycle noncontextual polytope**
  Mateus Araújo⋆, Marco Túlio Quintino⋆, Costantino Budroni, Marcelo Terra Cunha, Adán Cabello
  submitted – Phys. Rev. A 88, 022118 (2013)
  Section 3.5

⋆These authors contributed equally to this work.

Introduction

Can we predict the result of a flipped coin?

Our scientific belief suggests that if we know the material of the coin, how it was tossed, and all informations about the wind... we could, in principle, predict its result with certainty. But due to our ignorance on these variables, and the complexity of the system, we are forced to describe a flipping coin with probabilities.

As illustrated by one of Einstein's most famous quotes [5] "Nature doesn't know chance, it operates on mathematical principles. As I have said so many times, God doesn't play dice with the world", intrinsic probabilistic events causes discomfort in various physicists. Einstein was worried about the probabilities presented on the axioms of quantum mechanics, but at that time, he (and all other physicists) were not aware of one of the most astonishing properties of quantum mechanics, it predicts *nonlocal* correlations. The definition of nonlocal correlations, first published by John Bell in 1964, formalises the idea of physical scenarios in which results cannot, even in principle, be predicted with certainty.

Can we decide whether randomness in physical systems is really intrinsic? Are there real life experiments in which probabilities cannot be interpreted as an ignorance of some "hidden variable"? What are the consequences and limitations of this interpretation? The main motivation of this dissertation is to explore the necessary mathematical tools to understand these questions, and also, to answer some of them.

In chapter 1 we introduce what is known as a *Bell scenario* by studying collections of multivariable probability distributions. We define *multipartite boxes* and, using some vocabulary of convex geometry, we discuss the concepts of *non-signalisation*, *quantumness*, and *nonlocality*. Exploring these definitions, we prove the existence of quantum nonlocal boxes and discuss its consequences. We also present some explicit examples in the CHSH scenario, the simplest one that opens room for nonlocality.

In chapter 2 we present known examples, results, and proof techniques for multipartite boxes. We discuss some concepts of quantum mechanics which allows systems with nonlocal statistics. In order to gain some insights on multipartite boxes, we analyse the CHSH scenario more closely, and present some phenomena that only take place on more complex ones. We end this chapter by discussing one application of nonlocal boxes, cryptographic protocols in which the only assumption necessary for its security is that non-signalling boxes cannot exist. These proposals motivate a class of *device independent* protocols, which allows us to infer non-trivial properties of physical systems just by analysing its statistical data.

In chapter 3 we explore *general boxes* in what is known as a *contextuality scenario*, that generalises the idea of a Bell scenario to a wider collection of multivariable distributions. When dealing with general boxes, we overcome the interpretation of different parties pressing buttons, and discuss some simple and natural scenarios that were missed behind the multipartite assumption. We end this chapter by presenting a new result [4]: we analyse the n-cycle scenario and characterise its noncontextual polytope by its tight noncontextuality inequalities. This is the first time that a noncontextual polytope is fully described in a scenario with an arbitrary number of settings.

Since we presented and discussed various results on nonlocal/contextual boxes, in chapter 4 we seek for a *Bell test*, physical realisation of such boxes. We present some previous efforts and its *loopholes*, physical assumptions that may not be well justified. After proving some known bounds on efficiency requirements for attaining a loophole free Bell test, we explore them on an ideal physical setup that exhibit nonlocal statistical data. In this setup, we show that maximal quantum (CHSH) violation is attainable [1], and in some cases we have a CHSH violation even using photodetectors with arbitrarily low efficientcies [2]. We end the chapter by presenting a new proposal of an optical implementation of Bell test involving photodetection, homodyne, and atomic measurements [3].

Chapter 1

Correlations on N-partite boxes

God **does** play dice.

The main focus of this chapter is to introduce some vocabulary and basic concepts that are explored in the rest of the text. Although all these results and definitions are well known, we hope to clarify and organize then in a consistent notation with the concept of *multipartite boxes*.

1.1 Convex geometry

In Euclidean space, a set is convex if given any two points A, B, the line AB joining them lies entirely within that set. Intuitively, this means that the set is connected, in the sense that you connect a segment between any two points without leaving the set, and the set has no dents in its perimeter.

The notion of convex combination is also intimately close to the idea of probabilities: the results of a flipped coin may be understood as a convex combination between head and tails. Since convex geometry definitions are inspired in our everyday geometric experiences in low dimensions[1], various theorems sound very natural.

We remark that convex geometry has applications in many branches of science [6, 7, 8], but here we only provide a brief introduction on the necessary concepts for studying correlations on boxes. For a more complete starting point, we suggest the classic [9], the first chapter of [10], and the online lecture notes [11].

Definition 1 (Convex set)**.** *A set $\mathcal{C} \subseteq \mathbb{R}^n$ is convex if*

$$pc_1 + (1-p)c_2 \in \mathcal{C} \quad \forall c_1, c_2 \in \mathcal{C}, \text{ and } p \in [0,1].$$

The dimension of a convex set is the minimal dimension of the Euclidean space $\mathbb{R}^d$ necessary to describe $\mathcal{C}$.

[1] Usually, in $\mathbb{R}^2$ or $\mathbb{R}^3$.

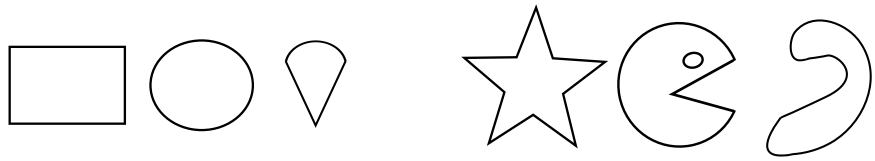

Figure 1.1: Three illustrations of convex and three illustrations of non-convex sets.

Definition 2 (Probability distribution)**.** *A distribution associated with K different results is a function* $p : \{i\}_{i=1}^{K} \to [0,1]$ *in which*

$$p_i \geq 0 \quad \textit{and} \sum_{i=1}^{K} p_i = 1.$$

An M-variable distribution is a function $p : \times_{m=1}^{M}\{i_m\} \to [0,1]$ *in which*

$$p_{i_1 i_2 \dots i_M} \geq 0 \quad \textit{and} \sum_{i_1 i_2 \dots i_M} p_{i_1 i_2 \dots i_M} = 1.$$

Definition 3 (Convex combination)**.** *A convex combination of elements* c_i *on a convex set* $\mathcal{C}$ *with distribution p is another element* $c \in \mathcal{C}$ *defined by*

$$c = \sum_i c_i p_i.$$

In this dissertation we will be interested in what is called a *convex body*, bounded closed convex sets with non-empty interior. In other words, convex bodies are the sets that are bounded by (possibly, infinitely) hyperplanes, where a hyperplane H is a (convex) set of points $\boldsymbol{x}$ that can be represented by a linear equation with real coefficients a_i, that is[2]

$$H = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{x}.\boldsymbol{a} = b\} \quad \boldsymbol{a} \in \mathbb{R}^n, \ b \in \mathbb{R}. \tag{1.1}$$

Please note that a hyperplane generalizes our notion of plane for any dimension: the equation $ax + by = 0$ in $\mathbb{R}^2$ represents a line that intercepts the origin and has an angular coefficient $-a/b$, the equation $z = 0$ in $\mathbb{R}^3$ represents the standard xy plane.

Convex bodies always contain some special points called *extremal points* or *pure points*, that cannot be written as non-trivial convex combinations of others[3]. Note that pure points always lie in the boundary of the convex set, but the boundary usually contains non-pure points as well. All non-pure points are called *mixed*.

Although we need all extreme points to describe the interior of a convex body as its convex combinations, a single point always can be written as a convex combination of $d+1$ extremal points, where d is the dimension of the convex set. This result is known as Carathéodory theorem and its proof can be found on wikipedia [12] or in any introductory convex geometry book.

[2]Here $\boldsymbol{x}.\boldsymbol{a}$ stands for the canonical inner product in $\mathbb{R}^n$.

[3]A convex combination is trivial if one of the probabilities p_i is equal to one.

Theorem 1 (Carathéodory)**.** *Any point of a convex body $\mathcal{C} \subseteq \mathbb{R}^n$ can be written as a convex combination of $d+1$ extreme points.*

A supporting hyperplane of a convex set $\mathcal{C}$ is a hyperplane that intersect $\mathcal{C}$ and is such that the all elements of $\mathcal{C}$ lies in one of the closed halfspaces formed by the hyperplane.

Definition 4 (Supporting hyperplane)**.** *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set, and $H := \{x \in \mathbb{R}^n | \boldsymbol{a}.\boldsymbol{x} = b\}$ be a hyperplane. H is a supporting hyperplane if H and $\mathcal{C}$ have at least one element in common and the halfspace $H_s := \{x \in \mathbb{R}^n | \boldsymbol{a}.\boldsymbol{x} \leq b\}$ contains $\mathcal{C}$.*

The convex hull of a set $\mathcal{A} \in \mathbb{R}^n$ is the smallest convex set that contains $\mathcal{A}$. For example the convex hull of a unit sphere $\mathcal{S}^{n-1} := \{\boldsymbol{x} \in \mathbb{R}^n | \ \|\boldsymbol{x}\| = 1\}$ is a unit ball $\mathcal{B}^n := \{\boldsymbol{x} \in \mathbb{R}^n | \ \|\boldsymbol{x}\| \leq 1\}$, the convex hull of the set of points $\mathcal{V} := \{(-1,-1), (-1,1), (1,-1), (1,1)\}$ is a square $\mathcal{S} := \{[-1,1] \times [-1,1]\}$.

We are now ready for the definition of *convex polytope*.

Definition 5 (Convex Polytope)**.** *A polytope[4] is a set that can be represented by the convex hull of a finite point subset of $\mathbb{R}^N$.*

Definition 6 (k-face)**.** *Let $\mathcal{C}$ be a convex set and $\mathcal{S}$ one of its supporting hyperplanes. The intersection $\mathcal{F} := \mathcal{S} \cap \mathcal{C}$ is a face of $\mathcal{C}$. A face of dimension k is called a k-face and has some special names for some specific values of k: 0-face is a vertex; 1-face is an edge; $n-1$ face is a facet.*

A set is written in $\mathcal{V}$-representation if its described only by its vertices. By definition, polytopes always admits a $\mathcal{V}$-representation. Our intuition suggests that convex polytopes can be described by its facets as well, that is, its tangent hyperplanes. This intuition is correct, and we call this halfspace description as $\mathcal{H}$-representation. This equivalence between these representations is ensured by the Weyl-Minkowisky theorem, and again, a proof can be found in basic convex geometry books.

Theorem 2 (Weyl-Minkowisky)**.** *All convex polytopes $P \subseteq \mathbb{R}^n$ admit a unique $\mathcal{H}$-representation with a finite number of linear inequalities.*

These two different representations have different applications, and suggest a natural question: how can we obtain the $\mathcal{H}$ ($\mathcal{V}$)-representation by knowing the vertex (facet) representation? This problem is known as the *vertex (facet) enumeration problem*, we'll see in section 1.6 that this problem is equivalent[5] to a central problem on nonlocality.

1.2 Single boxes

We now start our study of correlations, boxes, and probabilities. Before formal definitions, let's understand some motivation behind it.

[4]Since we are only concerned in convex scenario, we may use the word polytope and convex polytope interchangeably.

[5]More formally, they belong to the same complexity class.

Imagine an experimental setup constructed in a physics lab. Alice, that has no idea on how the experiment was constructed, knows all probabilities to obtain a certain result. To focus on the probabilistic description, we could imagine that the whole experimental setup is inside a *black box*[6], in which the only thing Alice can do is to press one of I different buttons, called *inputs*, to obtain one of O different *outputs*.

For example, consider a black box with 1 input and 2 outputs conveniently labelled by 0 and 1. This problem is fully described by a random variable that may assume values 0 and 1, and the statistical data of this black box is a probability distribution vector in $\mathbb{R}^2$,

$$\boldsymbol{p} = \begin{bmatrix} p_0 \\ p_1 \end{bmatrix}.$$

By trivial generalization we see that an O output single button black box can be described by an O-variable distribution vector in $\mathbb{R}^O$. Following the same line, a black box with I different inputs and O different outputs per input[7] is described by I different O-variable distributions, that can be represented as a vector $\mathbb{R}^{IO}$. This vector that completely describes the statistical data of a given experiment will be called a *single box*.

Adopting the convention[8] $p_{a|A_x}$ for the probability to obtain the outcome a after choosing the input A_x and labelling the inputs and outputs with natural numbers, a single black box with I inputs and O outputs can be described by

$$\boldsymbol{p} = \begin{bmatrix} p_{0|A_0} & \cdots & p_{0|A_{I-1}} \\ p_{1|A_0} & \cdots & p_{1|A_{I-1}} \\ \vdots & \ddots & \vdots \\ p_{O-1|A_0} & \cdots & p_{O-1|A_{I-1}} \end{bmatrix}. \tag{1.2}$$

Remark 1 (Representation of distributions)**.** *Maybe, the clearest notation for the probability of Alice to obtain the outcome a after choosing the input A_x would be $p(A_x = a|A_x)$, but in more general scenarios this notation easily becomes cumbersome. For the sake of compactness, we will always make the identification*

$$p(A_x = a|A_x) \equiv p_{a|A_x}. \tag{1.3}$$

Remark 2 (Box representations)**.** *Usually, vector space elements are represented by column (or row) matrix, but in this text we will always represent boxes using $n \times m$ matrices. For example, in a linear algebra text, the vector $\boldsymbol{p} \in \mathbb{R}^{IO}$ represented by an $I \times O$ matrix in equation (1.2) would probably appear as*

$$\boldsymbol{p} = \begin{bmatrix} p_{0|A_0} & p_{1|A_0} & \cdots & p_{O-1|A_0} & p_{0|A_1} & \cdots & p_{0|A_{I-1}} & \cdots & p_{1|A_{I-1}} & p_{O-1|A_{I-1}} \end{bmatrix}.$$

[6] The term black box, borrowed from computer scientists, is used to talk about objects without specifying its internal structure.

[7] We could also assume that each input has a different number of outputs, but this generalization has minor interest.

[8] This is an useful compact notation for studying various different distributions. For more details on this convention, please read remark 1.

Please note that our matrix representation is more compact, which is useful for high dimensional vectorial spaces. And in our context, we have the interesting property that all columns are probability distributions, that is, single boxes are stochastic matrices.

Definition 7 (Single-box). *Let $p_{a|A_x}$ be the probability of Alice to obtain the outcome a after choosing the input A_x. A single box with I inputs and O outputs is collection of distributions that can be represented by a vector $\boldsymbol{p} \in \mathbb{R}^{IO}$ where elements satisfy*

$$p_{a|A_x} \geq 0 \; \forall a, A_x, \qquad \sum_a p_{a|A_x} = 1 \quad \forall A_x.$$

The set of all single-boxes that can be constructed with I inputs and O outputs is called $\mathcal{B}(I, O)$.

It is also important to remark that a single box is just a convenient way to express the probabilities of I different random variables that can assume O different values.

We can easily check that $\mathcal{B}(I, O)$ is a convex set: convex combinations of two boxes $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ always satisfies the positivity and the normalisation condition:

$$\begin{aligned} \sum_a \Big(t p_{a|A_x} + (1-t)\tilde{p}_{a|A_x} \Big) &= t \sum_a \Big(p_{a|A_x} \Big) + (1-t) \sum_a \Big(\tilde{p}_{a|A_x} \Big) = \\ &= t + 1 - t \\ &= 1. \end{aligned}$$

And by noting that a point of this convex set is pure *iff* it represents a box with a deterministic probability distribution[9], we see that $\mathcal{B}(I, O)$ is a convex polytope.

1.3 Bipartite boxes

Imagine now Alice shares one black box experiment with Bob. That is, in a different laboratory, Bob has access to I_B different input buttons and may have O_B different output results. Now, we also assume that Alice and Bob only obtain their outputs after both press their input buttons[10]

A complete description of this bipartite scenario must specify all probabilities $p_{ab|A_x B_y}$ of Alice and Bob obtaining the results a and b when they press A_x and B_y. This can be made by specifying the probabilities of I_A and I_B, two variable distributions that can assume O_A, and O_B different values. As in a

[9] A probability distribution is deterministic if contains only zeros and ones.

[10] We also implicitly assume that the time order of pressing buttons does not affect the outputs, that is, Alice pressing x and then Bob pressing y is the same as Bob pressing y and then Alice pressing x. This assumption is related to the non-signalling condition, discussed on section 1.3.1.

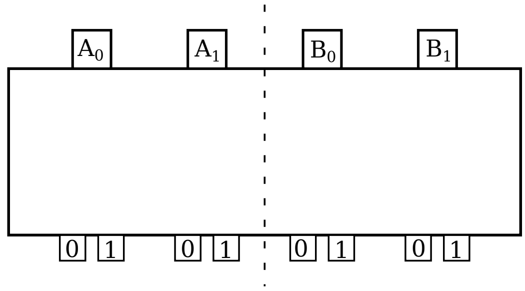

Figure 1.2: A box illustrating the scenario $(2, 2, 2)$, more discussed on section 1.8.

single box scenario, we represent this probabilities in a vector in $\mathbb{R}^{I_A I_B O_A O_B}$,

$$\boldsymbol{p} = \begin{bmatrix} p_{00|A_0B_0} & \cdots & p_{00|A_{I_A-1}B_{I_B-1}} \\ p_{01|A_0B_0} & \cdots & p_{01|A_{I_A-1}B_{I_B-1}} \\ \vdots & \ddots & \vdots \\ p_{O_A-1,O_B-1|A_0B_0} & \cdots & p_{O_A-1,O_B-1|A_{I_A-1}B_{I_B-1}} \end{bmatrix}. \tag{1.4}$$

Definition 8 (Bipartite scenario)**.** *Let $\mathcal{I}_A = \{A_x\}_1^{I_A}$, $\mathcal{O}_A = \{a\}_1^{O_A}$ be the set of inputs/outputs of Alice and $\mathcal{I}_B = \{B_y\}_1^{I_B}$, $\mathcal{O}_B = \{b\}_1^{O_B}$ be the set of inputs/outputs of Bob. The 4-tuple $\mathcal{S} = (\mathcal{I}_A, \mathcal{O}_A; \mathcal{I}_B, \mathcal{O}_B)$ is a bipartite box scenario.*

Since the label of inputs and outputs has no deep meaning, we also use the shorthand notation $\mathcal{S} = (I_A, O_A; I_B, O_B)$, and for the case where $I_A = I_B$ and $O_A = O_B$ we just write $\mathcal{S} = (2, I, O)$.

Definition 9 (Bipartite box)**.** *Let $p_{ab|xy}$ be the probability of Alice and Bob obtaining the outcomes a and b when they press A_x and B_y, respectively. A bipartite box in the scenario $\mathcal{S} = (I_A, O_A; I_B, O_B)$ is a collection of distributions that can be described by a vector $\boldsymbol{p} \in \mathbb{R}^d$ $(d = I_A I_B O_A O_B)$ in which elements satisfies*

$$p_{ab|A_xB_y} \geq 0 \;\forall a, b, A_x, B_y, \qquad \sum_{a,b} p_{ab|A_xB_y} = 1 \qquad \forall \quad A_x, B_y. \tag{1.5}$$

The set of all bipartite boxes that can be constructed in a specific scenario is $\mathcal{B}(\mathcal{S})$. For the case where $I_A = I_B$ and $O_A = O_B$ we just write $\mathcal{B}(2, I, O)$.

As in the single box scenario, bipartite boxes are just one convenient way to write a set of two variable probability distributions. And the set $\mathcal{B}(I_A, O_A; I_B, O_B)$ is the convex hull of all bipartite deterministic boxes.

The structure of two variable distributions naturally suggest the notion of correlation between two parts. But before exploring correlations, it is useful to define *marginal probabilities* to connect the idea of two-variable with single-variable distributions. Marginal probability formalizes what one would call the probability of Alice to obtain a certain value a when A_x and B_y are pressed.

Definition 10 (Marginal probability). *The marginal probabilities $p_{a|A_x;B_y}$ of a two variable distribution with coefficients $p_{ab|A_xB_y}$ is*

$$p_{a|A_x;B_y} := \sum_b p_{ab|A_xB_y}.$$

We also define the marginal probabilities $p_{b|B_y;A_x}$ as

$$p_{b|B_y;A_x} := \sum_a p_{ab|A_xB_y}.$$

Remark 3. *Maybe, the clearest notation for the probability of Alice and Bob to obtain the outcomes a and b when they press A_x and B_y would be $p(A_x = a, B_y = b|A_x, B_y)$. So the marginals would be defined as*

$$p(A_x = a|A_x, B_y) := \sum_b p(A_x = a, B_y = b|A_x, B_y);$$
$$p(B_y = b|A_x, B_y) := \sum_a p(A_x = a, B_y = b|A_x, B_y).$$

But for the sake of compactness we will always make the identifications

$$p(A_x = a, B_y = b|A_x, B_y) \equiv p_{ab|A_xB_y};$$
$$p(A_x = a|A_x; B_y) \equiv p_{a|A_x;B_y};$$
$$p(B_y = b|A_x; B_y) \equiv p_{a|B_y;A_x}.$$

Intuitively, a two variable distribution is uncorrelated if we can describe it only by its marginals. More formally, a two variable distribution $p_{\cdot\,\cdot|A_xB_y}$ is completely uncorrelated if, for each fixed x, y,

$$p_{ab|A_xB_y} = p_{a|A_x;B_y} p_{b|B_y;A_x} \; \forall a, b. \tag{1.6}$$

It is interesting to note that deterministic two variable distributions are always uncorrelated, to prove this we just need to use the fact that deterministic distributions only have events of probability zero or one. So randomness is a necessary condition for correlations between two variable distributions.

In order to gain some interpretation on correlations it is interesting to invoke a simple result from probability theory.

Theorem 3. *For any fixed x, y, a two variable distribution $p_{\cdot\,\cdot|A_xB_y}$ that can assume O_A different values on the first variable and O_B different values on the second variable can be written as*

$$p_{ab|A_xB_y} = \sum_\lambda \pi_{xy}(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda} \tag{1.7}$$

for some distribution $\pi : \Lambda_{xy} \to [0, 1]$ and one variable distributions $p_{\cdot|A_x;\lambda}$, $p_{\cdot|B_y;\lambda}$.

Moreover, we can always chose the one variable distributions $p_{\cdot|A_x;\lambda}$ and $p_{\cdot|B_y;\lambda}$ to be deterministic, and the cardinality of the set Λ_{xy} is bounded by $O_AO_B + 1$.

Proof. The set of all two variable distributions that can assume O_A different values on the first variable and O_B different values on the second variable is a convex polytope with $O_A O_B$ deterministic distributions as vertices. Now we use deterministic $p_{.|A;\lambda}$ and $p_{.|B;\lambda}$ to construct the vertices of the two variables distribution polytope via the product of marginals $p_{ab|A_xB_y;\lambda} := p_{a|A;\lambda}p_{b|B;\lambda}$. Since all points inside a polytope can be written as convex combinations of its vertices, all two variable distributions can be written in the form of $p_{ab|A_xB_y} = \sum_\lambda \pi_{xy}(\lambda)p_{ab|A_xB_y;\lambda}$. We we guarantee the existence of a set Λ_{xy} with cardinality $O_A O_B + 1$ by invoking Carathéodory theorem. □

The equation (1.7) allows the interpretation:

> In principle, all two-variable distributions are uncorrelated and deterministic, probabilities arise as an ignorance on λ.

Due to this interpretation, λ is usually referred as a *hidden variable*, since we are forced to work with probabilities because we cannot have access to its value.

Since bipartite boxes are just a collection of two variable distributions, theorem 3 has two natural extensions for bipartite boxes.

Corollary 1. *All probabilities of bipartite boxes admit the representation*

$$p_{ab|A_xB_y} = \sum_\lambda \pi_{xy}(\lambda)p_{a|A_x;\lambda}p_{b|B_y;\lambda} \tag{1.8}$$

for some distributions $\pi_{xy} : \Lambda_{xy} \to [0,1]$, *and one variable distributions* $p_{.|A_x;\lambda}$ *and* $p_{.|B_y;\lambda}$.

Corollary 2. *All probabilities of bipartite boxes admit the representation*

$$p_{ab|A_xB_y} = \sum_\lambda \pi(\lambda)p_{a|A_x;B_y\lambda}p_{b|B_y;A_x\lambda} \tag{1.9}$$

for some distribution $\pi : \Lambda \to [0,1]$, *and one variable distributions* $p_{.|A_x;B_y\lambda}$ *and* $p_{.|B_y;A_x\lambda}$.

Proof. To prove this corollary we just need that, since the distributions of the vertices of the bipartite box polytope $\mathcal{B}(2, I, O)$ are deterministic, they can always be written as

$$p_{ab|A_xB_y} = p_{a|A_x;B_y\lambda}p_{b|B_y;A_x\lambda},$$

by setting $p_{a|A_x;B_y\lambda}$ and $p_{b|B_y;A_x\lambda}$ as deterministic. Now, it follows by convexity that all probabilities of bipartite boxes can be written in the form of 1.9. □

Another simple result that will be explored in next sections is that $\mathcal{B}(I_A, O_A) \otimes \mathcal{B}(I_B, O_B)$ has the same dimension of $\mathcal{B}(I_A, O_A; I_B, O_B)$, this implies that we can always represent a bipartite box $\boldsymbol{p}_{AB} \in B(2, I, O)$ as linear combinations of tensor products of single boxes. The reader who is familiar with quantum mechanics can now start making a parallel between bipartite boxes and composed quantum systems.

1.3.1 Non-signalling

Consider a two output, two input, deterministic bipartite box defined by

$$p_{00|A_0B_0} = 1, \quad p_{10|A_0B_1} = 1, \quad p_{01|A_1B_0} = 1, \quad p_{11|A_1B_1} = 1. \tag{1.10}$$

Other equivalent (and possibly, more enlightening) way to define this box is writing[11]

$$a = y, b = x,$$

that is, Alice's output is Bob's input, and *vice versa*. Can we ask questions like "What is the probability of Alice to obtain 0 while she presses A_0?"?

From marginal probabilities, we know that the probability of Alice obtaining 0 while she presses A_0 depends on Bob's choice B_y. If $y = 0$, Alice obtains 0, if $y = 1$, she obtains 1. So, if $p_{0|A_0;B_0} \neq p_{0|A_0;B_1}$, Alice's probability to obtain 0 as pressing A_0, $p_{0|A_0}$, may not have a precise meaning.

A natural question is, when can we talk about $p_{0|A_0}$? Well, if Alice's marginal probabilities satisfy $p_{0|A_0;B_0} = p_{0|A_0;B_1}$, it is fair to define $p_{0|A_0} := p_{0|A_0;B_0} = p_{0|A;0B_1}$. This condition is called *non-signalling*.

Definition 11 (Non-Signalling)**.** *A bipartite box is non-signalling if its probabilities satisfy*[12]

$$\begin{aligned} \sum_b p_{ab|A_xB_y} &= \sum_b p_{ab|A_xB_{y'}} := p_{a|A_x} \qquad \forall a, A_x, B_y, B_{y'} \\ \sum_a p_{ab|A_xB_y} &= \sum_a p_{ab|A_{x'}B_y} := p_{b|B_y} \qquad \forall b, A_x, A_{x'}, B_y \end{aligned} \tag{1.11}$$

The set of all non-signalling bipartite boxes is denoted by $\mathcal{NS}(I_A, O_A; I_B, O_B)$, or just $\mathcal{NS}(2, I, O)$ in case $I_A = I_B$ and $O_A = O_B$.

Remark 4. *Using the clear notation $p(A_x = a, B_y = b|A_x, B_y)$ suggested in remark 3 we see that the non-signalling condition allows us suppress the indexes $|A_x$, $|B_y$ in our notation. That is, without ambiguity, we can write*

$$\begin{aligned} p(A_x = a) &:= p(A_x = a|A_x) =: p(A_x = a|A_x, B_y); \\ p(B_y = b) &:= p(B_y = c|B_y) =: p(B_y = b|A_x, B_y); \\ p(A_x = a, B_y = b) &=: p(A_x = a, B_y = b|A_x, B_y). \end{aligned}$$

In the non-signalling scenario, we can understand A_x and B_y as usual random variables. We will see in chapter 3 that the existence of the probabilities $p(A_x = a)$ and $p(B_y = b)$ are the conditions imposed in a what we call a marginal scenario.

Informally, non-signalling is a necessary and sufficient condition for talking about two "individual" parts[13], if Alice presses the input button A_x, she can obtain an output a_x independently of weather Bob pressed a button or not[14].

[11] In fact, all deterministic boxes admit this input/output representation.

[12] We remark that some authors refer tor non-signalling bipartite boxes as *behaviours*. This term was coined in one of Tsirelson's seminal papers [13].

[13] In signalling boxes, we have no rights to talk about individual parts, because one side can "manipulate" the other probabilities by exploring $p_{a|A_x;B_y} \neq p_{a|A_x;B_{y'}}$.

[14] In information theory [14] this means that this box cannot be used for communication. That is, exist a protocol in which Alice and Bob can use the inputs and outputs to send messages.

It is very important to remark that the non-signalling condition does not forbid correlations, we can still have $p_{ab|A_xB_y} \neq p_{a|A_x}p_{b|B_y}$. One good example is the *perfect correlated*, that in the box in the scenario where Alice and Bob have two outputs, can be defined as[15]

$$p_{00|A_xB_y} = p_{11|A_xB_y} = 1/2.$$

We can check that all distributions of this boxes are correlated, $p_{ab|A_xB_y} \neq p_{a|A_x}p_{b|B_y}$, and also the non-signalling condition holds true:

$$\begin{aligned} p_{00|A_xB_y} + p_{01|A_xB_y} &= p_{00|A_xB'_y} + p_{01|A_xB'_y} = 1/2 \quad \forall x, y, y' \,; \\ p_{00|A_xB_y} + p_{10|A_xB_y} &= p_{00|A'_xB'_y} + p_{10|A'_xB'_y} = 1/2 \quad \forall x, x', y \,. \end{aligned}$$

Another important remark is that the linearity of equations[16] (1.11) ensures that $\mathcal{NS}(I_A, O_A; I_B, O_B)$ is a polytope[17].

1.3.2 Local

For those who know λ, there are no correlations.

Theorem 3 ensures the classical intuition that all correlations between probabilities arise as a possible ignorance from an inaccessible information. What can we say about box correlations?

Before exploring box correlations, it is important to have a good comprehension of distribution correlations. Suppose Eve prepares a bipartite single input box[18], $\boldsymbol{p} \in \mathcal{B}(2, 1, 2)$ using the following rule: she flips a coin; if heads, she outputs 0 for Alice and 1 for Bob; if tails she does the opposite. Assuming that the only source of uncertainty in this experiment is Eve's coin's result, and that the coin returns heads with probability[19] $p(h)$, without knowing the result of Eve's coin, Alice and Bob's best description of this box is

$$p_{ab} = p(h)p_{a|A;h}p_{b|B;h} + p(t)p_{a|A;t}p_{b|B;t}, \tag{1.12}$$

or more explicitly

$$\begin{aligned} p_{00} &= p(h)p_{0|A;h}p_{0|B;h} + p(t)p_{0|A;t}p_{0|B;t} = 0; \\ p_{01} &= p(h)p_{0|A;h}p_{1|B;h} + p(t)p_{0|A;t}p_{1|B;t} = p(h); \\ p_{10} &= p(h)p_{1|A;h}p_{0|B;h} + p(t)p_{1|A;t}p_{0|B;t} = p(t); \\ p_{11} &= p(h)p_{1|A;h}p_{1|B;h} + p(t)p_{1|A;t}p_{1|B;t} = 0. \end{aligned} \tag{1.13}$$

---

[15] Also note that the normalisation condition that implies that $p_{01|A_xB_y} = p_{10|A_xB_y} = 0$

[16] For real numbers, $a = b \iff a \geq b$, and $a \leq b$.

[17] Intersections of a polytope with a finite number of hyperplanes is also a polytope.

[18] Although single input bipartite boxes are just two-variable probability distribution, we chose this notation for further generalization.

[19] And tails with probability $p(t) = 1 - p(h)$.

These equations say that Alice and Bob cannot predict their outcome with certainty only because they do not have access to the coin's result. In principle, Eve (and/or others) can describe the whole experience by $p_{01|h} = 1$ and $p_{10|t} = 1$. In fact, we will later prove that Eve can construct (and then foresee the results with certainty) all single input bipartite boxes just by flipping coins.

What can we say about general bipartite boxes? From corollary 2 we can write $p_{ab|A_xB_y} = \sum_\lambda \pi_{xy}(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda}$, but we need various different $\pi_{..|xy}$ distributions that depend on Alice and Bob's input choice. This means that in the general case, Eve must know their inputs to construct a bipartite box. One could ask: what class of boxes can Eve construct if she cannot have access to Alice and Bob's inputs A_x and B_y?

Definition 12 (Locality). *A bipartite box is local if all probabilities can be written as*

$$p_{ab|A_xB_y} = \sum_\lambda \pi(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda} \qquad \forall a, x, y, y', \tag{1.14}$$

for some distribution $\pi : \Lambda \to [0,1]$ and single variable distributions $p_{A_x;\lambda}$ and $p_{B_y;\lambda}$.

The set of all local bipartite boxes is $\mathcal{L}(I_A, O_A; I_B, O_B)$, or just $\mathcal{L}(2, I, O)$ in case $I_A = I_B$ and $O_A = O_B$.

This definition was first presented in John Bell's 1964 seminal paper [15]. Bell was motivated by studying correlations that have a same common cause: the distribution π over λ.

It must be clear that, for local boxes, we can always define Alice's marginal probabilities regardless of Bob's choice. That is, all local boxes are non-signalling.

Theorem 4. *$\mathcal{L}(I_A, O_A; I_B, O_B) \subseteq \mathcal{NS}(I_A, O_A; I_B, O_B)$.*

Proof. Since the role of A_x and B_y is symmetric in both definitions, we just need to prove that if we assume equation (1.14), the marginal $p_{a|A_x;B_y}$ does not depend on B_y. Explicitly:

$$\begin{aligned} p_{a|A_x;B_y} &:= \sum_b p_{ab|A_xB_y} \\ &= \sum_{b\lambda} \pi(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda}; \\ &= \sum_\lambda \pi(\lambda) p_{a|A_x;\lambda} (\sum_b p_{b|B_y;\lambda}); \\ &= \sum_\lambda \pi(\lambda) p_{a|A_x;\lambda}. \end{aligned}$$

□

For some very particular scenarios, locality is equivalent to non-signalling. In these cases we say that the scenario is trivial.

Theorem 5 ([20]).

$$\mathcal{B}(2,1,O) = \mathcal{L}(2,1,O);$$
$$\mathcal{L}(I_A,O_A;I_B,1) = \mathcal{NS}(I_A,O_A;I_B,1);$$
$$\mathcal{L}(1,O_A;I_B,O_B) = \mathcal{NS}(1,O_A;I_B,O_B).$$

This theorem follows as a corollary of theorem 22, to be proved on chapter 3. In fact, this theorem is even stronger. These particular scenarios are the only ones in which non-signalling implies locality. It may be surprising to know that there are nonlocal boxes that respect the non-signalling condition, in section 1.8 we will illustrate this fact in the $(2,2,2)$ scenario.

We now exhibit an equivalent definition for local boxes that explore the vectorial character of box spaces.

Theorem 6. *A bipartite box* $\boldsymbol{p} \in \mathcal{B}(I_A,O_A;I_B,O_B)$ *is local* iff *it can be written as*

$$\boldsymbol{p} = \sum_\lambda \pi(\lambda)\boldsymbol{p}_A^\lambda \otimes \boldsymbol{p}_B^\lambda, \tag{1.15}$$

where $\pi : \Lambda \to [0,1]$ *is a distribution,* $\boldsymbol{p}_A^\lambda \in \mathcal{B}(I_A,O_A)$*, and* $\boldsymbol{p}_B^\lambda \in \mathcal{B}(I_B,O_B)$*.*

Proof. To prove this theorem we just need to recognize that the components of $\boldsymbol{p}$ are exactly

$$p_{ab|A_xB_y} = \sum_\lambda \pi(\lambda)p_{a|A_x;\lambda}p_{b|B_y;\lambda}.$$

□

This simple theorem provides an interesting parallel between nonlocality and *quantum entanglement*. In section 2.1 we define quantum entanglement and present a brief discussion on this analogy.

Before ending this section, we make an important connection between deterministic non-signalling boxes and local boxes.

Theorem 7. *A bipartite box is local* iff *is a convex combination of bipartite non-signalling deterministic boxes.*

Proof. As proved before, all local boxes are non-signalling, so we only need to prove that convex combinations of non-signalling deterministic boxes are local.

Since any deterministic two variable distribution is completely uncorrelated, any deterministic non-signalling box has the probabilities in the form

$$p_{ab|A_xB_y;\lambda} = p_{a|A_x;\lambda}p_{b|B_y;\lambda}.$$

Now, we just check that convex combinations of these probabilities is exactly the definition of locality,

$$\sum_\lambda \pi(\lambda)p_{ab|A_xB_y;\lambda} = \sum_\lambda \pi(\lambda)p_{a|A_x;\lambda}p_{b|B_y;\lambda}.$$

□

[20] It may be useful to recall that $\mathcal{S} = (N,I,O)$ represents a scenario with N parties, I inputs per party, and O outputs per input, and $\mathcal{S}' = (I_A,O_A;I_B,O_B)$ represents a bipartite scenario where Alice has I_A inputs and O_A outputs per input, Bob has I_B inputs and O_B outputs per input.

Corollary 3. *The set of local boxes is a convex polytope in which vertices are deterministic non-signalling boxes.*

Due to this fact, the set of all bipartite local boxes is also known as the *local polytope*.

1.3.3 Quantum

How can we describe boxes that represent quantum experiments? In other perspective, what class of boxes can Eve prepare with quantum systems? The ones who are familiarised with quantum mechanics may say: "we just need to prepare a certain quantum state and choose some measurement operators, then the probabilities to obtain a certain output will be given by the quantum measurement postulate". And that is exactly the definition of quantum boxes, the ones that can be constructed by measurements on quantum systems.

We will now define all quantum objects that are necessary for understanding quantum boxes, but nothing more than that. Readers that do not feel comfortable with quantum mechanics are invited to read Nielsen and Chuang's book [16], or for introductions to quantum mechanics more focused on nonlocality we suggest [17, 18, 19, 20].

Definition 13 (Quantum Theory)**.** *Let $\mathcal{H}_n$ be an n-dimensional Hilbert space over $\mathbb{C}$. A quantum state is a trace* 1 *positive semidefinite*[21] *linear operator $\rho : \mathcal{H}_n \to \mathcal{H}_n$.*

A set of operators $\mathcal{M} = \{M^i\}$ is a quantum measurement set if all $M^i : \mathcal{H}_n \to \mathcal{H}_n$ are positive semidefinite operators that sumf to identity, $\sum_i M^i = I$.

The probability to obtain the result i after subjecting the state ρ to the quantum measurement $\mathcal{M}$ is $p_{i|\rho\mathcal{M}} = \operatorname{tr} \rho M^i$.

We are now able to define bipartite quantum boxes.

Definition 14 (Quantum Boxes)**.** *A bipartite box is quantum if each probability can be written as*

$$p_{ab|A_xB_y} = \operatorname{tr}(\rho_{AB} A_x^a \otimes B_y^b) \qquad \forall a, b, x, y,$$

for a certain quantum state $\rho_{AB} : \mathcal{H}_n \otimes \mathcal{H}_m \to \mathcal{H}_n \otimes \mathcal{H}_m$ and quantum measurement sets $\mathcal{A}_x = \{A_x^i\}$, $A_x^i : \mathcal{H}_n \to \mathcal{H}_n$ and $\mathcal{B}_y = \{B_y^j\}$,$B_y^j : \mathcal{H}_m \to \mathcal{H}_m$.

The set of all quantum bipartite boxes is called $\mathcal{Q}(I_A, O_A; I_B, O_B)$, or just $\mathcal{Q}(2, I, O)$ in case $I_A = I_B$ and $O_A = O_B$.

Please note that this definition does not specify the dimension of the Hilbert space associated to the quantum system, that is, to understand $\mathcal{Q}(I_A, O_A; I_B, O_B)$ we may need to deal with infinite dimensional vector spaces[22].

Quantum measurements on one single part[23] cannot alter the results on the other one. In our text, this means that bipartite quantum boxes are non-signalling.

[21] An operator $\rho : \mathcal{H}_n \to \mathcal{H}_n$ is positive semidefinite if satisfies $\langle v, \rho v\rangle \geq 0$, $\forall v \in \mathcal{H}_n$.

[22] In fact, in [21] the authors suggest that some quantum boxes need measurements on infinite dimensions to be constructed.

[23] They are also called *local measurements*, and are represented by measurement operators that take the form $\{M_n \otimes I\}$.

Theorem 8. $\mathcal{Q}(I_A, O_A; I_B, O_B) \subseteq \mathcal{NS}(I_A, O_A; I_B, O_B)$.

Proof. The proof follows from a straightforward calculation that checks that all boxes $\boldsymbol{p} \in \mathcal{Q}(I_A, O_A; I_B, O_B)$ satisfy the conditions $p_{a|A_xB_y} = p_{a|A_xB'_y}$, and $p_{b|A_xB_y} = p_{b|A'_xB_y}$:

$$\begin{aligned} p_{a|A_xB_y} &= \sum_b p_{ab|A_xB_y}; \\ &= \sum_b \operatorname{tr}(\rho_{AB} A_x^a \otimes B_y^b); \\ &= \operatorname{tr}(\rho_{AB} A_x^a \otimes (\sum_b B_y^b)); \\ &= \operatorname{tr}(\rho_{AB} A_x^a \otimes (I)). \end{aligned}$$

□

Also, all local boxes can be simulated by quantum systems. Before proving that fact, we would like to explicitly show that all probability distributions can be simulated by measurements on quantum systems.

Theorem 9. *All multivariable probability distributions p can be simulated by performing measurements on quantum systems.*

Proof. For simplicity, we will first assume that p is a single variable distribution represented by probabilities $\{p_i\}_{i=1}^K$.

Define quantum a state $\rho : \mathcal{H}_K \to \mathcal{H}_K$ and quantum measurement operators $M^j : \mathcal{H}_K \to \mathcal{H}_K$

$$\rho = \sum_i^K p_i \Pi_i, \quad M^j = \Pi_j,$$

where $\{\Pi_i\}_1^K$ is a set of unidimensional orthogonal projectors. Now, by straightforward calculation we can check that these operators define a valid quantum system[24] and $\operatorname{tr} \rho M^i = p_i$.

For multivariable distributions, we can use exactly the same technique, we just need to enlarge our Hilbert space. □

Theorem 10. $\mathcal{L}(I_A, O_A; I_B, O_B) \subseteq \mathcal{Q}(I_A, O_A; I_B, O_B)$.

Proof. Just note that it is possible to construct a quantum state and some measurement operators to obtain any local box probabilities. For example, we can always use the state

$$\rho_{AB} = \sum_\lambda \pi(\lambda) \rho_A^\lambda \otimes \rho_B^\lambda$$

and measurement operators that satisfy (see theorem 9)

$$\operatorname{tr}(\rho_A^\lambda A_a^x) = p_{a|A_x;\lambda}, \quad \text{and} \quad \operatorname{tr}(\rho_B^\lambda B_b^y) = p_{b|B_y;\lambda},$$

[24] $\rho, M^i \geq 0, \quad \operatorname{tr} \rho = 1, \quad \sum_i M^i = I.$

to obtain

$$\mathrm{tr}\,(\rho_{AB} A_a^x \otimes B_b^y) = \sum_\lambda \pi(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda}.$$

□

Although all local boxes are quantum, there are nonlocal quantum boxes. This result is known as Bell's theorem[25] [15], and has some deep consequences on foundations of quantum mechanics. The first immediately consequence is the absence of generalization of theorem 7: some quantum systems cannot be described by convex combinations of deterministic boxes.

The set $\mathcal{Q}(I_A, O_A; I_B, O_B)$ is convex [22, 23], but differently from previously discussed box sets, it is not a polytope[26]. This fact makes the structure of quantum box sets much more complex.

Theorem 11. *The set $\mathcal{Q}(I_A, O_A; I_B, O_B)$ is convex.*

Proof. We need to show that if $\boldsymbol{p}^1, \boldsymbol{p}^2 \in \mathcal{Q}(I_A, O_A; I_B, O_B)$, then $\lambda \boldsymbol{p}^1 + (1-\lambda)\boldsymbol{p}^2 \in \mathcal{Q}(I_A, O_A; I_B, O_B)$.

If $\boldsymbol{p}^1, \boldsymbol{p}^2 \in \mathcal{Q}(I_A, O_A; I_B, O_B)$, there are quantum states $\rho^1, \rho^2 : \mathcal{H}_n \otimes \mathcal{H}_n \to \mathcal{H}_n \otimes \mathcal{H}_n$ and quantum measurements operators $A_x^{i,a}, B_y^{j,b} : \mathcal{H}_n \to \mathcal{H}_n$ such that

$$\begin{aligned} p^1_{ab|A_x B_y} &= \mathrm{tr}\,(\rho^1 A_x^{1,a} \otimes B_y^{1,b}); \\ p^2_{ab|A_x B_y} &= \mathrm{tr}\,(\rho^2 A_x^{2,a} \otimes B_y^{2,b}). \end{aligned}$$

Now, we define a state $\tilde{\rho} : \mathcal{H}_{2n} \otimes \mathcal{H}_{2n} \to \mathcal{H}_{2n} \otimes \mathcal{H}_{2n}$ and measurement operators $\tilde{A}_x^a, \tilde{B}_y^b : \mathcal{H}_{2n} \to \mathcal{H}_{2n}$ as

$$\begin{aligned} \tilde{\rho} &:= \lambda \rho^1 \otimes \Pi^1 \otimes \Pi^1 + (1-\lambda)\rho^2 \otimes \Pi^2 \otimes \Pi^2; \\ \tilde{A}_x^a &:= A_x^{1,a} \otimes \Pi^1 + A_x^{2,a} \otimes \Pi^2; \\ \tilde{B}_x^a &:= B_y^{1,b} \otimes \Pi^1 + B_y^{2,b} \otimes \Pi^2, \end{aligned}$$

where $\Pi^1, \Pi^2 : \mathcal{H}_2 \to \mathcal{H}_2$ are orthogonal unidimensional projectors. Now we check that the probabilities of this new quantum system are exactly the convex combination of the older ones,

$$\begin{aligned} &\tilde{p}_{ab|\tilde{A}_x \tilde{B}_y} = \mathrm{tr}(\tilde{\rho} \tilde{A}_x^a \otimes \tilde{B}_y^b) = \\ &= \lambda\, \mathrm{tr}\,((\rho^1 \otimes \Pi^1 \otimes \Pi^1)[(A_x^{1,a} \otimes \Pi^1 + A_x^{2,a} \otimes \Pi^2) \otimes (B_y^{1,b} \otimes \Pi^1 + B_y^{2,b} \otimes \Pi^2)]) \\ &+ (1-\lambda)\, \mathrm{tr}\,((\rho^2 \otimes \Pi^2 \otimes \Pi^2)[(B_x^{1,a} \otimes \Pi^1 + B_x^{2,a} \otimes \Pi^2) \otimes (B_y^{1,b} \otimes \Pi^1 + B_y^{2,b} \otimes \Pi^2)]) \\ &= \lambda\, \mathrm{tr}\,(\rho^1 A_x^{1,a} \otimes B_y^{1,b}) + (1-\lambda)\, \mathrm{tr}\,(\rho^2 A_x^{2,a} \otimes B_y^{2,b}) \\ &= \lambda p^1_{ab|A_x B_y} + (1-\lambda) p^2_{ab|A_x B_y}. \end{aligned}$$

So we can always construct a quantum box that provides us the probabilities of the convex combination $\lambda \boldsymbol{p}^1 + (1-\lambda)\boldsymbol{p}^2$. □

[25] Some times also stated in phrases like "Quantum mechanics is a nonlocal theory", or "Quantum mechanics cannot be described by local hidden variables".

[26] In section 1.8 we will see that $\mathcal{Q}(I_A, O_A; I_B, O_B)$ has infinitely many extremal points.

Please note that we explored the freedom on the vector space dimension to prove the convexity of $\mathcal{Q}(I_A, O_A; I_B, O_B)$. In fact, if we require quantum states to lie in a Hilbert space with a fixed dimension, the set of quantum boxes may not be convex [24].

1.4 General properties on bipartite box sets

Before starting with multipartite boxes, we summarize some important properties of bipartite box sets. We also anticipate that these properties still hold for the multipartite scenario.

- They respect the hierarchical relation

  $$\mathcal{L}(\mathcal{S}) \subset \mathcal{Q}(\mathcal{S}) \subset \mathcal{NS}(\mathcal{S}) \subset \mathcal{B}(\mathcal{S}),$$

  where, $\mathcal{S} = (I_A, O_A; I_B, O_B)$ is a bipartite scenario with more than one output and input per part[27].

- All previous bipartite box sets are convex.

- Except from the quantum, all sets are polytopes.

1.5 Multipartite scenario

All previous concepts developed for bipartite boxes will be now generalized for an N-partite scenario. As N grows, the number of possibilities on signalization and nonlocality became larger. And although most generalizations are quite natural, some difficulties may appear due to a necessarily heavy notation. In order to help the readers to visualise multipartite boxes, we will provide illuminating examples.

1.5.1 Multipartite boxes

In the bipartite scenario, we allowed each part to have different number of outputs and inputs. This freedom has minor interest and may lead us into very cumbersome notation. So, for practical reasons, we will now assume that each part has access to I different inputs and O different outcomes per input.

Definition 15 (Multipartite scenario)**.** *Let $\mathcal{I}_n = \{i_n\}_{i=1}^{I}$, $\mathcal{O}_n = \{o_n\}_{o=1}^{O}$ be the set of inputs/outputs of the party n. The 3-tuple $\mathcal{S} = (N, \{\mathcal{I}_n\}^N, \{\mathcal{O}_n\}^N)$ is a multipartite box scenario.*

Since the label of inputs and outputs has no deep meaning, we also use the shorthand notation $\mathcal{S} = (N, I, O)$.

Definition 16 (Multipartite box)**.** *Let*

$$p_{\boldsymbol{o}|\boldsymbol{i}} := p_{o_1 o_2 \ldots o_N | 1_{i_1} 2_{i_2} \ldots N_{i_N}}$$

[27] If the parties have only one input or output, the sets may be equivalent.

be the probability of the parties $1, 2, \dots, N$ *to obtain the outcomes* $o_1, o_2, \dots, o_N$, *after they chose (respectively) the inputs* $i_1, i_2, \dots, i_N$. *An N-partite box in the scenario* (N, I, O) *is a collection of distributions that can be described as a vector* $\boldsymbol{p} \in \mathbb{R}^d$ $(d = (IO)^N)$ *in which elements satisfy*

$$p_{\boldsymbol{o}|\boldsymbol{i}} \geq 0 \quad \forall \boldsymbol{o}, \boldsymbol{i}, \qquad \sum_{\boldsymbol{o}} p_{\boldsymbol{o}|\boldsymbol{i}} = 1 \quad \forall \boldsymbol{i}. \tag{1.16}$$

The set of all N-partite boxes that can be described with I inputs and O outputs is $\mathcal{B}(N, I, O)$.

As in the bipartite scenario, the set $\mathcal{B}(N, I, O)$ is the convex hull of deterministic multipartite boxes, and its elements are just a convenient way of representing I^N different N-variable distributions.

The definition of marginal probabilities for N-variable distributions is completely analogous to the two variable ones, except by the difference that more variables allow to talk about more marginal possibilities. Let us take a closer look at the tripartite case.

The number $p_{o_1 o_2 o_3 | 1_{i_1} 2_{i_2} 3_{i_3}}$ represents the probability of parts 1, 2, 3 to obtain the outputs o_1, o_2, o_3, after pressing i_1, i_2, i_3. The distribution $p_{1_{i_1} 2_{i_2} 3_{i_3}}$ can provide us 6 different marginal probabilities:

$$\begin{aligned}
p_{o_1|1_{i_1};2_{i_2}3_{i_3}} &:= \sum_{o_2 o_3} p_{o_1 o_2 o_3|1_{i_1}2_{i_2}3_{i_3}}; \\
p_{o_2|2_{i_2};1_{i_1}3_{i_3}} &:= \sum_{o_1 o_3} p_{o_1 o_2 o_3|1_{i_1}2_{i_2}3_{i_3}}; \\
p_{o_3|3_{i_3};1_{i_1}2_{i_2}} &:= \sum_{o_1 o_2} p_{o_1 o_2 o_3|1_{i_1}2_{i_2}3_{i_3}}; \\
p_{o_1 o_2|1_{i_1}2_{i_2};3_{i_3}} &:= \sum_{o_3} p_{o_1 o_2 o_3|1_{i_1}2_{i_2}3_{i_3}}; \\
p_{o_2 o_3|2_{i_2}3_{i_3};1_{i_1}} &:= \sum_{o_1} p_{o_1 o_2 o_3|1_{i_1}2_{i_2}3_{i_3}}; \\
p_{o_1 o_3|1_{i_1}3_{i_3};2_{i_2}} &:= \sum_{o_2} p_{o_1 o_2 o_3|1_{i_1}2_{i_2}3_{i_3}}.
\end{aligned}$$

Before moving to the general definition, we remark some possible notation issues.

Remark 5. *As in* $p_{\boldsymbol{o}|\boldsymbol{i}} = p_{o_1 o_2 \dots o_N | 1_{i_1} 2_{i_2} \dots N_{i_N}}$, *we are going to use bold letters for representing a list of variables.*

- $\boldsymbol{l}$ *will represent parties involved, for example* $\boldsymbol{l} = 2, 3, 5$ *represents the parties* 2, 3, *and* 5.
- $\bar{\boldsymbol{l}}$ *is the complement of* $\boldsymbol{l}$, *for example if* $\boldsymbol{l} = 2, 3, 5$, *then* $\bar{\boldsymbol{l}} = 1, 4, 6, \dots, N$.
- *The set* $\mathcal{O}_{\boldsymbol{l}}$ *is the set of all possible outcomes related to the list* $\boldsymbol{l}$.

Definition 17 (Marginal probability). *Given an N variable distribution with coefficients* $p_{\boldsymbol{o}|\boldsymbol{i}}$ *and a list* $\boldsymbol{l}$ *we define the marginals as*

$$p_{\boldsymbol{o_l}|\boldsymbol{i_l};\boldsymbol{i_{\bar{l}}}} := \sum_{\mathcal{O}_{\bar{\boldsymbol{l}}}} p_{\boldsymbol{o_l}\boldsymbol{o_{\bar{l}}}|\boldsymbol{i_l}\boldsymbol{i_{\bar{l}}}}. \tag{1.17}$$

1.5.2 Non-signalling boxes

Definition 18 (Non-Signalling)**.** *An N-partite box $\boldsymbol{p} \in \mathcal{B}(N, I, O)$ is non-signalling if the marginal probabilities for all list of parties $\boldsymbol{l}$ do not depend on the other parties $\bar{\boldsymbol{l}}$ input choices. That is, all marginals can be written as $p_{\boldsymbol{o_l}|\boldsymbol{i_l};\boldsymbol{i_{\bar{l}}}} = p_{\boldsymbol{o_l}|\boldsymbol{i_l}}$.*

The set of all non-signalling boxes with N parties, I inputs per party, and O outputs per input is denoted by $\mathcal{NS}(N, I, O)$.

One interesting theorem proved in [19] states that if all marginals with $N-1$ variables are non-signalling, the whole box is NS. For example, in a tripartite scenario, if 1 cannot signal to 2 and 3 and that 2 cannot signal to 1 and 3, we deduce

$$\begin{aligned} \sum_{o_1,o_2} p_{o_1o_2o_3|i_1i_2i_3} &= \sum_{o_1,o_2} p_{o_1o_2o_3|i_1'i_2i_3} \quad \forall\, o_3, i_1'i_1i_2i_3 \\ &= \sum_{o_1,o_2} p_{o_1o_2o_3|i_1'i_2'i_3} \quad \forall\, o_3, i_1'i_1i_2'i_2i_3, \end{aligned} \tag{1.18}$$

which is the condition that 1 and 2 cannot signal to 3.

Another important remark is that, like in bipartite scenarios, $\mathcal{NS}(N, I, O)$ is a convex polytope and a multipartite box can be used for communication *iff* is signalling.

1.5.3 Local boxes

Definition 19 (Locality)**.** *An N-partite box is local if all probabilities can be written as*

$$p_{\boldsymbol{o}|\boldsymbol{i}} = \sum_\lambda \pi(\lambda) p_{o_1|1_{i_1};\lambda} p_{o_2|2_{i_2};\lambda} \cdots p_{o_N|N_{i_N};\lambda} \qquad \forall \boldsymbol{o}, \boldsymbol{i}, \tag{1.19}$$

for some distribution $\pi : \Lambda \to [0, 1]$ and single variable distributions $p_{x_{i_x};\lambda}$.

The set of all local multipartite boxes with N parties, I inputs per party, and O outputs per input is denoted by $\mathcal{L}(N, I, O)$.

We also remark that $\mathcal{L}(N, I, O)$ is the convex hull of all non-signalling deterministic boxes, and it is also called as *local polytope*. Also, the parallel with quantum entangled states still remains: a multipartite box is nonlocal *iff* it cannot be written as convex combination of tensor product of single boxes, $\boldsymbol{p} = \sum_\lambda \pi(\lambda) \boldsymbol{p}_1^\lambda \otimes \boldsymbol{p}_2^\lambda \otimes \ldots \boldsymbol{p}_N^\lambda$. The proof of these theorems are completely analogous to the ones proved for bipartite boxes.

An interesting question related to nonlocal multipartite boxes is: do we attain different kinds of nonlocality by exploring multipartite boxes? For example, in the tripartite case, are there nonlocal boxes that cannot be written as nonlocal bipartite boxes? This property, referred as *genuine multipartite nonlocality* may be trickier than one could imagine. For example, although we have important results on genuine multipartite nonlocality since 1987 [25, 26, 27], a "good" definition only became public in 2011 [28].

1.5.4 Quantum boxes

The boxes that we can construct with quantum systems follow from simple generalization. Also all other properties presented for the bipartite scenario still remain: quantum multipartite boxes form a convex set with infinitely many extremal points and respect $\mathcal{L}(N, I, O) \subseteq \mathcal{Q}(N, I, O) \subseteq \mathcal{NS}(N, I, O)$.

Definition 20 (Quantum boxes)**.** *An N-partite box is quantum if each probability can be written as*

$$p_{\boldsymbol{o}|\boldsymbol{i}} = \mathrm{tr}\,(\rho\, 1_{1_i}^{o_1} \otimes 2_{2_i}^{o_2} \otimes \ldots \otimes N_{N_i}^{o_N}) \qquad \forall \boldsymbol{o}, \boldsymbol{i},$$

for a certain quantum state $\rho : \mathcal{H}_n^{\otimes N} \to \mathcal{H}_n^{\otimes N}$ *and quantum measurement sets* $1_x = \{1_{1_i}^{o_1}\}$, $1_{1_i}^{o_1} : \mathcal{H}_n \to \mathcal{H}_n$, $2_x = \{2_{2_i}^{o_2}\}$, $2_{2_i}^{o_2} : \mathcal{H}_n \to \mathcal{H}_n$, $\ldots$, $N_x = \{N_{N_i}^{o_N}\}$, $N_{N_i}^{o_N} : \mathcal{H}_n \to \mathcal{H}_n$.

The set of all quantum N-partite boxes with I inputs per party and O outputs per input is $\mathcal{Q}(N, I, O)$.

1.5.5 General properties of multipartite box sets

As stated in section 1.4, multipartite boxes also obey some general properties.

- They respect the hierarchical relation

$$\mathcal{L}(\mathcal{S}) \subset \mathcal{Q}(\mathcal{S}) \subset \mathcal{NS}(\mathcal{S}) \subset \mathcal{B}(\mathcal{S}),$$

  where, $\mathcal{S} = (N, I, O)$ is a multipartite scenario with[28] $I, O \geq 2$.

- All previous multipartite box sets are convex.

- Except from the quantum, all sets are polytopes.

1.6 Bell inequalities and the facet enumeration problem

> When satisfied they indicate that the data may have, when not satisfied they indicate that the data cannot have, resulted from actual observation [29].
>
> George Boole

Given all probabilities of a certain box, how can we decide if this box is local, in the sense of definition (19). Geometrically, given $\boldsymbol{p} \in \mathcal{B}(N, I, O)$, how to decide whenever $\boldsymbol{p} \in \mathcal{L}(N, I, O)$?

The problem of deciding if a certain $\boldsymbol{p}$ lies inside a polytope is famous in computational geometry. One way to solve it is to check if $\boldsymbol{p}$ satisfies all halfspace inequalities. Theorem 7 guarantees that the vertices of the local polytope are the non-signalling deterministic boxes. So, we can reduce our

[28] If the parties have only one input or output, the sets may be equivalent.

question to the facet enumeration problem, that is, to find the $\mathcal{H}$-representation of a convex polytope.

There is an algorithm based on the Fourier-Motzkin elimination[29] that transforms the $\mathcal{V}$-representation into the $\mathcal{H}$-representation [32]. Moreover, there is a algorithm implemented in C that performs this transformation [33], so the problem of finding all facet inequalities is, in some sense, solved. We said in some sense because the time computational complexity of the best known algorithm for solving this problem is[30] $O(d!V^d)$ [35], where V stands for the vertex number and d is the dimension of the polytope[31].

In complexity theory, algorithms that have exponential complexity are non-efficient[32]. Readers that are not familiar with computation may take the following rule of thumb: "Non-efficient algorithms takes a very long time to be solved even with very good computers". If the best known algorithm to solve a certain problem is non-efficient, this problem is said to be computationally hard [36].

1.6.1 The facets of the local polytope

The first ideas on exploring correlation sets as convex polytopes were proposed by Froissart in 1981[37]. It took 12 years until someone (Boris Tsirelson) explored these geometric aspects on nonlocality [13]. But now, almost all papers on nonlocality use this geometric vocabulary.

Due to John Bell's seminal paper on non-locality [15], the hyperplanes that describe the local polytope are called *Bell inequalities*.

Definition 21 (Bell inequalities)**.** *The linear inequalities that are respected by all local boxes are called Bell inequalities.*

Inequalities that are equivalent to the positivity, normalisation, or non-signalling condition (equations (1.5) *and* (1.11)*), are said to be trivial.*

Bell inequalities defining facets are called tight.

We could not end this section without commenting the very curious fact that pioneer papers on nonlocality were published without mentioning that Bell inequalities actually date back to 1862, due to George Boole's works [29, 38]. Most famous for his contributions to mathematical logic, Boole also has some papers on probability in which he considered the question: How can we decide if a certain set of probabilities came from a valid statistical data. Immersed in a classical scenario, Boole could not think that a nonlocal set of probabilities might be physically possible, and after finding his inequalities he wrote:

---

[29] The first register to this algorithm dates back to 1826, in a work by Fourier [30]. Without knowing Fourier results, Motzkin rediscovered it 1936 in his thesis [31].

[30] We also remark the existence of a discussion on how should we measure the complexity of this problem [23, 34].

[31] In section 1.10 we will see that the dimension of the polytope grows exponentially in the number of parts. So the complexity of our main problem is even larger.

[32] An algorithm is efficient if, in the worse case, its time-complexity has a polynomial dependence on the problem's input.

> When satisfied they indicate that the data may have, when not satisfied they indicate that the data cannot have, resulted from actual observation.

Readers who are interested on Boole's motivation may look for Pitowsky's review [39]. We also remark that Pitowsky's book on quantum logic [22] has a similar approach to nonlocality using ideas similar to the ones proposed by Boole.

1.7 Full-correlation sets

In this section we will develop a sufficient[33] method to tell if a certain box is nonlocal. This method consists in defining a projector map that can be used to infer some properties of correlation boxes without specifying all its probabilities $p_{\boldsymbol{o}|\boldsymbol{i}}$.

1.7.1 Full-correlator for bipartite dichotomic boxes

For a two variable distribution $p_{\cdot\cdot|A_xB_y}$ that can assume the values 0 and 1, we define the correlator map as

$$\begin{aligned}\langle A_xB_y\rangle &:= p_{00|A_xB_y} + p_{11|A_xB_y} - p_{01|A_xB_y} - p_{10|A_xB_y} \\ &= p_{a=b|A_xB_y} - p_{a\neq b|A_xB_y}.\end{aligned} \tag{1.20}$$

The correlator of a distribution satisfies some interesting properties: if Alice and Bob outputs are always the same, $\langle A_xB_y\rangle = 1$; if they are always the opposite, $\langle A_xB_y\rangle = -1$; if they are completely random, $\langle A_xB_y\rangle = 0$.

This notation is inspired by the fact that if we treat A_x and B_y as random variables that can assume the values ±1, the correlator is formally equivalent to the expected value of the random variable A_xB_y,

$$\begin{aligned}\langle A_xB_y\rangle :=& p_{+1+1|A_xB_y} + p_{-1-1|A_xB_y} - p_{-1+1|A_xB_y} - p_{+1-1|A_xB_y} \\ =& p_{a=b|A_xB_y} - p_{a\neq b|A_xB_y}.\end{aligned}$$

The correlator map for distributions has a natural extension for dichotomic bipartite boxes. The full-correlator box map $FC : \mathbb{R}^{(IO)^2} \to \mathbb{R}^{I^2}$ applied in a box $\boldsymbol{p}$ is a full-correlation box $\boldsymbol{p}_F = FC(\boldsymbol{p})$ that lists all possible correlators. For example, full-correlation box of

$$\boldsymbol{p} = \begin{bmatrix} p_{00|A_0B_0} & \cdots & p_{00|A_{I_A-1}B_{I_B-1}} \\ p_{01|A_0B_0} & \cdots & p_{01|A_{I_A-1}B_{I_B-1}} \\ p_{10|A_0B_0} & \cdots & p_{10|A_{I_A-1}B_{I_B-1}} \\ p_{11|A_0B_0} & \cdots & p_{11|A_{I_A-1}B_{I_B-1}} \end{bmatrix}$$

is just

$$FC(\boldsymbol{p}) = \left[\langle A_0B_0\rangle, \quad \langle A_0B_1\rangle, \quad \ldots, \quad \langle A_{I_A-1}B_{I_B-1}\rangle\right].$$

[33] But not necessary.

It must also be clear that various different multipartite boxes may lead to the same full-correlation one. That is, the correlation map is not a bijection.

This map is interesting because $FC(\boldsymbol{p})$ is usually much more simple to deal than $\boldsymbol{p}$, and we are going to prove that if $FC(\boldsymbol{p})$ is nonlocal, then $\boldsymbol{p}$ is nonlocal.

Inspired by the full-correlator of two-variable dichotomic distributions, we also define the single correlator of a distribution $p_{.|A_x}$ as

$$\langle A_x \rangle := p_{0|A_x} - p_{1|A_x}. \tag{1.21}$$

In next sections, we will see that we can explore single correlator map to simplify some proofs and notations. Please also note that if A_x is a random variable that can assume values $+1$ and -1, we can understand the single correlator as the expected value of A_x.

1.7.2 Full-correlator for multipartite dichotomic boxes

In the general case, the correlator map is a function that attributes a number $C \in [-1, 1]$ to an N-variable distribution that can assume 2 different values.

Definition 22 (Correlator for distributions)**.** *Let $p_{o_1 o_2 \ldots o_N | 1_{i_1} 2_{i_2} \ldots N_{i_N}}$ be the probabilities of a dichotomic N-variable distribution. We set the dichotomic values to be $+1$ and -1 and define the N-correlator map as the expected value*

$$\langle 1_{i_1} 2_{i_2} \ldots N_{i_N} \rangle := \sum_{\mathcal{O}} o_1 o_2 \ldots o_N \; p_{o_1 o_2 \ldots o_N | 1_{i_1} 2_{i_2} \ldots N_{i_N}}.$$

An equivalent approach is to set the dichotomic values to be 0 and 1 and define the correlator map as

$$\langle 1_{i_1} 2_{i_2} \ldots N_{i_N} \rangle := p_{o_1 \oplus o_2 \oplus \ldots \oplus o_N = 0 | 1_{i_1} 2_{i_2} \ldots N_{i_N}} - p_{o_1 \oplus o_2 \oplus \ldots \oplus o_N = 1 | 1_{i_1} 2_{i_2} \ldots N_{i_N}},$$

where $\oplus$ stands for sum modulo 2.

Definition 23 (Full-correlator for boxes)**.** *Let $\boldsymbol{p} \in \mathcal{B}(N, I, 2)$ be an N-partite box. The full-correlation N-partite box $FC(\boldsymbol{p})$ is the list of all I^N O-variable correlators associated to $\boldsymbol{p}$.*

We are now able to define the full-correlation version of all previous box sets $\mathcal{B}$, $\mathcal{NS}$, $\mathcal{L}$, $\mathcal{Q}$ as all full-correlation boxes that can be constructed with them.

Definition 24. *The full-correlation box set $\mathcal{FB}(N, I, 2)$ is the set generated by all full-correlation boxes of $\mathcal{B}(N, I, 2)$. That is,*

$$\boldsymbol{p}_F \in \mathcal{FB}(N, I, 2) \;\; \textit{if} \;\; \exists \boldsymbol{p} \in \mathcal{B}(N, I, 2), \;\; \textit{such that } FC(\boldsymbol{p}) = \boldsymbol{p}_F.$$

We also define $\mathcal{FNS}(N, I, 2)$, $\mathcal{FL}(N, I, 2)$, and $\mathcal{FQ}(N, I, 2)$ in a completely analogous way.

Note that the set all full-correlation multipartite boxes $\mathcal{FB}(N, I, 2)$ is an hypercube of dimension I^N, since there are no restriction on the probabilities, its vertices are vectors with entries ± 1. We will now show that is also the case for full-correlation non-signalling boxes[34].

Theorem 12. $\mathcal{FB}(N, I, 2) = \mathcal{FNS}(N, I, 2)$, *moreover these sets are hypercubes of dimension* I^N *with vertices being vectors with entries* ± 1.

Proof. Clearly, $\mathcal{FNS}(N, I, 2) \subseteq \mathcal{FB}(N, I, 2)$, so we just need to prove that the vertices $(\pm 1, \pm 1, \dots, \pm 1)$ are attainable with non-signalling full-correlation boxes.

Define two N-variable dichotomic distributions $p^{\pm} : \{-1, +1\}^{\times N} \to [0, 1]$ as

$$p^{\pm}_{o_1 o_2 \dots o_N} := \frac{|\pm 1 + o_1 o_2 \dots o_N|}{2^N},$$

which are positive by definition and sum to unity,

$$\begin{aligned}
\sum_{o_1 o_2 \dots o_N} p^{\pm}_{o_1 o_2 \dots o_N} &= \sum_{o_1 o_2 \dots o_N} \frac{|\pm 1 + o_1 o_2 \dots o_N|}{2^N};\\
&= \sum_{o_1 o_2 \dots o_{N-1}} \left(\frac{|\pm 1 + o_1 o_2 \dots o_{N-1}|}{2^N} + \frac{|\pm 1 - o_1 o_2 \dots o_{N-1}|}{2^N} \right);\\
&= \sum_{o_1 o_2 \dots o_{N-1}}^{\prod_i o_i = 1} \left(\frac{|\pm 1 + 1| + |\pm 1 - 1|}{2^N} \right) + \sum_{o_1 o_2 \dots o_{N-1}}^{\prod_i o_i = -1} \left(\frac{|\pm 1 - 1| + |\pm 1 + 1|}{2^N} \right);\\
&= \left(2^{(N-1)} \frac{2}{2^N} \right);\\
&= 1,
\end{aligned}$$

where $\sum_{o_1 o_2 \dots o_{N-1}}^{\prod_i o_i = \pm 1}$ stands for summing for all o_i such that the product $\prod_i o_i$ is equal to ± 1.

From the above calculation we can see that all marginals of $p^{\pm}$ are uniform. For example:

$$\begin{aligned}
p^{\pm}_{o_1 o_2 \dots o_{N-1}} &= \sum_{o_N} p^{\pm}_{o_1 o_2 \dots o_N}\\
&= \frac{|\pm 1 + o_1 o_2 \dots o_{N-1}|}{2^N} + \frac{|\pm 1 - o_1 o_2 \dots o_{N-1}|}{2^N}\\
&= \frac{2}{2^N},
\end{aligned}$$

and

$$\begin{aligned}
p^{\pm}_{o_1} &= \sum_{o_2 \dots o_N} p^{\pm}_{o_1 o_2 \dots o_N};\\
&= \sum_{o_2 \dots o_{N-1}} \left(\frac{|\pm 1 + o_1 o_2 \dots o_{N-1}|}{2^N} + \frac{|\pm 1 - o_1 o_2 \dots o_{N-1}|}{2^N} \right);\\
&= \frac{1}{2}.
\end{aligned}$$

[34] I acknowledge Gláucia Murta for presenting me this nice and simple proof.

Now note the correlators associated to the distributions $p^{\pm}$ are equal to ± 1

$$\sum_{o_1 o_2 \dots o_N} p^{\pm}_{o_1 o_2 \dots o_N} = \sum_{o_1 o_2 \dots o_N} o_1 o_2 \dots o_N \frac{|\pm 1 + o_1 o_2 \dots o_N|}{2^N};$$
$$= \sum_{o_1 o_2 \dots o_{N-1}} \left(\frac{o_1 o_2 \dots o_N \; |\pm 1 + o_1 o_2 \dots o_{N-1}|}{2^N} - \frac{o_1 o_2 \dots o_N \; |\pm 1 - o_1 o_2 \dots o_{N-1}|}{2^N} \right);$$
$$= \sum_{o_1 o_2 \dots o_{N-1}}^{\prod_i o_i = 1} \left(\frac{|\pm 1 + 1|}{2^N} \right) - \sum_{o_1 o_2 \dots o_{N-1}}^{\prod_i o_i = -1} \left(\frac{|\pm 1 - 1|}{2^N} \right);$$
$$= \pm 1.$$

Since all marginals of $p^{\pm}$ are uniform distributions, boxes that are constructed with distributions $p^{\pm}$ are always non-signalling.

Now note that we have a technique to generate all vertices of the hypercube $\mathcal{FB}(N, I, 2)$ by applying the full-correlation map into non-signalling boxes. We just need to associate the distribution $p^{\pm}$ to the inputs in which its correlator is[35] ± 1.

□

We may use the full-correlation version of box sets to learn some properties of the non-full ones. For example, we can use $FC(\boldsymbol{p})$ to infer nonlocality or "non-quantumness".

Theorem 13.

$$FC(\boldsymbol{p}) \notin \mathcal{FL}(N, I, 2) \implies \boldsymbol{p} \notin \mathcal{L}(N, I, 2)$$
$$FC(\boldsymbol{p}) \notin \mathcal{FQ}(N, I, 2) \implies \boldsymbol{p} \notin \mathcal{Q}(N, I, 2)$$

Proof. The counterpart of the theorem is

$$\boldsymbol{p} \in \mathcal{L}(N, I, 2) \implies FC(\boldsymbol{p}) \in \mathcal{FL}(N, I, 2);$$
$$\boldsymbol{p} \in \mathcal{Q}(N, I, 2) \implies FC(\boldsymbol{p}) \in \mathcal{FQ}(N, I, 2),$$

which is true by definition. □

The converse of this theorem is not true, but since full-correlation sets are much simpler than the non-full ones, this theorem is extremely useful.

1.8 CHSH: an explicit example

As theorem 5 states, the set $\mathcal{B}(2, 2, 2)$ is the simplest one in which we can find nonlocal boxes that are not signalling. In this section we provide a concrete example of a local polytope, nonlocal boxes, and Bell inequalities. With this example we develop some basic techniques that can be used on more complex

[35] For example, if we want that the full-correlator $\langle A_x B_y C_z \rangle$ to be ± 1, we just set $p_{\dots|A_x B_y C_z} := p^{\pm}$.

scenarios. We also point out what are the particular features of $\mathcal{B}(2,2,2)$ and what characteristics hold for more general multipartite boxes.

Due to Clauser, Horn, Shimony, and Holt's 1969 seminal paper [40], the scenario $(2,2,2)$ is known as the CHSH scenario. One interesting point is that the inequalities presented by them were the first tight Bell inequalities ever published. Bell did use some inequalities in his 1964 seminal paper [15], he also proved that we can construct quantum boxes that are nonlocal, but he was only concerned with one specific kind of correlations and his inequality did not represent a facet of the local polytope.

We warn the readers that the approach used in this section is different from the ones used in more traditional books of quantum mechanics [41, 16]. Usually they present one CHSH inequality (see subsection 1.8.2), prove that it is respected by all local boxes, and exhibit an example of a quantum boxes that violates it. Here we will discuss and analyse all previously defined box sets individually.

1.8.1 $\mathcal{B}(2,2,2)$

First, we establish the convention that all box probabilities will be represented in the form

$$\boldsymbol{p} = \begin{bmatrix} p_{00|A_0B_0} & p_{00|A_0B_1} & p_{00|A_1B_0} & p_{00|A_1B_1} \\ p_{01|A_0B_0} & p_{01|A_0B_1} & p_{01|A_1B_0} & p_{01|A_1B_1} \\ p_{10|A_0B_0} & p_{10|A_0B_1} & p_{10|A_1B_0} & p_{10|A_1B_1} \\ p_{11|A_0B_0} & p_{11|A_0B_1} & p_{11|A_1B_0} & p_{11|A_1B_1} \end{bmatrix}. \tag{1.22}$$

In this convention, the deterministic box in which Alice and Bob always obtain the output 0 independently of their input choice is

$$\boldsymbol{p}_{00|A_xB_y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1.23}$$

Now note that the vertices of the polytope $\mathcal{B}(2,2,2)$ are vectors like (1.23), matrices in which columns represents deterministic probability distributions. So, $\mathcal{B}(2,2,2)$ has 4^4 vertices, since each column can assume one of four different deterministic distributions.

The $\mathcal{H}$-representation of $\mathcal{B}(2,2,2)$ is given by all normalisation and positivity inequalities:

$$\sum_{ab} p_{ab|A_xB_y} = 1,\ \forall A_x, B_y, \qquad p_{ab|A_xB_y} \geq 0,\ \forall a, b, A_x, B_y. \tag{1.24}$$

Using the normalisation condition of distributions, we can prove that we need 3×4 parameters to describe an element of $\mathcal{B}(2,2,2)$. So, $\mathcal{B}(2,2,2)$ is a convex polytope of dimension 12 with 256 vertices.

1.8.2 $\mathcal{L}(2,2,2)$

With theorem 7, it is easy to decide which vertices of $\mathcal{B}(2,2,2)$ are vertices of $\mathcal{L}(2,2,2)$, we just need to check which ones satisfy the non-signalling condition. We also recall that in the non-signalling scenario, it is always possible to understand A_x and B_y as random variables that can assume two different values. This suggests a useful notation for the vertices of the local polytope: the vector $\boldsymbol{L}_{a_0 a_1 b_0 b_1}$ represents the deterministic box in which we always have $A_x = a_x$ and $B_y = b_y$. More explicitly,

$$\boldsymbol{L}_{0000} = \begin{bmatrix} 1&1&1&1 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{0001} = \begin{bmatrix} 1&0&1&0 \\ 0&1&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{0010} = \begin{bmatrix} 0&1&0&1 \\ 1&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{0011} = \begin{bmatrix} 0&0&0&0 \\ 1&1&1&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix};$$

$$\boldsymbol{L}_{0100} = \begin{bmatrix} 1&1&0&0 \\ 0&0&0&0 \\ 0&0&1&1 \\ 0&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{0101} = \begin{bmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{bmatrix},\ \boldsymbol{L}_{0110} = \begin{bmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{bmatrix},\ \boldsymbol{L}_{0111} = \begin{bmatrix} 0&0&0&0 \\ 1&1&0&0 \\ 0&0&0&0 \\ 0&0&1&1 \end{bmatrix};$$

$$\boldsymbol{L}_{1000} = \begin{bmatrix} 0&0&1&1 \\ 0&0&0&0 \\ 1&1&0&0 \\ 0&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{1001} = \begin{bmatrix} 0&0&1&0 \\ 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \end{bmatrix},\ \boldsymbol{L}_{1010} = \begin{bmatrix} 0&0&0&1 \\ 0&0&1&0 \\ 0&1&0&0 \\ 1&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{1011} = \begin{bmatrix} 0&0&0&0 \\ 0&0&1&1 \\ 0&0&0&0 \\ 1&1&0&0 \end{bmatrix};$$

$$\boldsymbol{L}_{1100} = \begin{bmatrix} 0&0&1&1 \\ 0&0&0&0 \\ 1&1&0&0 \\ 0&0&0&0 \end{bmatrix},\ \boldsymbol{L}_{1101} = \begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 1&0&1&0 \\ 0&1&0&1 \end{bmatrix},\ \boldsymbol{L}_{1110} = \begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&1&0&1 \\ 1&0&1&0 \end{bmatrix},\ \boldsymbol{L}_{1111} = \begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 1&1&1&1 \end{bmatrix}.$$

Please note that this technique is completely general, and can be used to generate the local vertices of all multipartite boxes.

The $\mathcal{H}$-representation of $\mathcal{L}(2,2,2)$ is given by trivial Bell inequalities (equations (1.24)) and the so called *CHSH inequalities*[36]:

$$-2 \leq -\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle + \langle A_1 B_1 \rangle \leq 2; \tag{1.25a}$$
$$-2 \leq +\langle A_0 B_0 \rangle - \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle + \langle A_1 B_1 \rangle \leq 2; \tag{1.25b}$$
$$-2 \leq +\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_0 \rangle + \langle A_1 B_1 \rangle \leq 2; \tag{1.25c}$$
$$-2 \leq +\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2, \tag{1.25d}$$

that together are necessary and sufficient conditions [42, 22] for locality in the scenario $(2,2,2)$.

Although it may be hard to prove that these are in fact all non-trivial Bell inequalities for this scenario, it is easy to check that all local boxes respect them. Just recall that the vertices of the local polytope are sets of deterministic (so, uncorrelated) distributions that satisfy the non-signalling condition. For such boxes, we can write

$$\langle A_i B_j \rangle = \langle A_i \rangle \langle B_j \rangle, \tag{1.26}$$

where the correlators can assume the values $+1$ or -1.

Now, using equation (1.26), we see that when analysing vertices of the local polytope, we can always write

$$\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle = \langle A_0 \rangle (\langle B_0 \rangle + \langle B_1 \rangle) + \langle A_1 \rangle (\langle B_0 \rangle - \langle B_1 \rangle)$$

and check that the right hand side can only assume the values $+2$ or -2, and by convexity we see that all local boxes respect the CHSH inequalities.

[36] In chapter 3 we will prove that these are all facet inequalities. But remember that we can always find all Bell inequalities by using computer algorithms [22].

1.8.3 $\mathcal{NS}(2,2,2)$

Clearly, deterministic non-signalling boxes are vertices of the non-signalling polytope. We also know that all other deterministic boxes are not members of[37] $\mathcal{NS}(2,2,2)$. So, how could one find the vertices of $\mathcal{NS}(2,2,2)$?

We know the $\mathcal{H}$-representation of non-signalling polytopes (non-signalling, positivity, and normalisation inequalities), but in some cases it is interesting to have the $\mathcal{V}$-representation. For the polytope $\mathcal{NS}(2,2,O)$, Barrett *et al* proposed an interesting systematic technique to find all vertices [43]. But in general, the problem of finding vertices of the non-signalling polytope is computationally equivalent to the one of finding Bell inequalities[38] [44].

Due to Popescu and Rohrlich's 1994 result that we can violate the CHSH inequality up to its algebraic maximum with non-signalling boxes [45], the nonlocal vertices of $\mathcal{NS}(2,2,2)$ are known as PR-boxes. A curious fact is that Popescu and Rohrlich (and apparently, various others) did not know that these boxes had been previously discovered by two different researchers (also independently) Rastall [46] and Tsirelson [47, 13] both in 1985.

The 8 PR-boxes are:

$$\boldsymbol{PR}_1 = \frac{1}{2}\begin{bmatrix} 1&1&1&0\\ 0&0&0&1\\ 0&0&0&1\\ 1&1&1&0 \end{bmatrix},\ \boldsymbol{PR}_2 = \frac{1}{2}\begin{bmatrix} 1&1&0&1\\ 0&0&1&0\\ 0&0&1&0\\ 1&1&0&1 \end{bmatrix},\ \boldsymbol{PR}_3 = \frac{1}{2}\begin{bmatrix} 1&0&1&1\\ 0&1&0&0\\ 0&1&0&0\\ 1&0&1&1 \end{bmatrix},\ \boldsymbol{PR}_4 = \frac{1}{2}\begin{bmatrix} 0&1&1&1\\ 1&0&0&0\\ 1&0&0&0\\ 0&1&1&1 \end{bmatrix};$$
$$\boldsymbol{PR}_5 = \frac{1}{2}\begin{bmatrix} 0&0&0&1\\ 1&1&1&0\\ 1&1&1&0\\ 0&0&0&1 \end{bmatrix},\ \boldsymbol{PR}_6 = \frac{1}{2}\begin{bmatrix} 0&0&1&0\\ 1&1&0&1\\ 1&1&0&1\\ 0&0&1&0 \end{bmatrix},\ \boldsymbol{PR}_7 = \frac{1}{2}\begin{bmatrix} 0&1&0&0\\ 1&0&1&1\\ 1&0&1&1\\ 0&1&0&0 \end{bmatrix},\ \boldsymbol{PR}_8 = \frac{1}{2}\begin{bmatrix} 1&0&0&0\\ 0&1&1&1\\ 0&1&1&1\\ 1&0&0&0 \end{bmatrix}.$$

We can also check that the PR-boxes violate the CHSH inequalities up to 4 (-4), its algebraic maximum (minimum).

1.8.4 $\mathcal{Q}(2,2,2)$

Even for this simple scenario, it is not known how all extreme points or the inequalities that describe quantum box sets. This problem is so hard that in 1993, Tsirelson asked [13]

"Does the set of quantum behaviours[39] this admit a description by a finite number of analytic inequalities? Or even – polynomial inequalities?", a question tof which we have no answer.

Although we do not have the complete description of $\mathcal{Q}(2,2,2)$, we have some important partial results. For example, it is known that all quantum

[37] These other boxes are signalling.

[38] Algorithms that are used to find $\mathcal{V}$-representation can also be used to find the $\mathcal{H}$-representation [44].

[39] Instead of quantum boxes, he used the term *quantum behaviours*, that are exactly the same objetcs.

boxes must obey[40] the *Tsirelson bound* [48], that is

$$
\begin{aligned}
-2\sqrt{2} &\leq -\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle + \langle A_1B_1\rangle \leq 2\sqrt{2};\\
-2\sqrt{2} &\leq +\langle A_0B_0\rangle - \langle A_0B_1\rangle + \langle A_1B_0\rangle + \langle A_1B_1\rangle \leq 2\sqrt{2};\\
-2\sqrt{2} &\leq +\langle A_0B_0\rangle + \langle A_0B_1\rangle - \langle A_1B_0\rangle + \langle A_1B_1\rangle \leq 2\sqrt{2};\\
-2\sqrt{2} &\leq +\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle \leq 2\sqrt{2}. \qquad (1.27)
\end{aligned}
$$

These inequalities are proved to be tangent halfspaces on $\mathcal{Q}(2,2,2)$, that is, we can always find a state and some quantum measurements to saturate these inequalities [48].

It is also interesting to remark that, since the PR-boxes (nonlocal vertices of the non-signalling polytope) violate one CHSH inequality up to ± 4, the Tsirelson bound implies that they cannot be constructed by performing measurements on quantum systems.

In [49, 50], Navascues, Pironio, and Acín introduced an infinite hierarchy of necessary conditions for any set of quantum boxes. Exploring the first step of the proposed hierarchy[41], they proved that all elements of $\mathcal{Q}(2,2,2)$ must satisfy

$$
-\pi \leq asinD_{00} + asinD_{01} + asinD_{10} - asinD_{11} \leq \pi, \qquad (1.28)
$$

where *asin* is the inverse of the sine function and

$$
D_{xy} := \langle A_xB_y\rangle - \frac{\langle A_x\rangle\langle B_y\rangle}{\sqrt{\left(1-\langle A_x\rangle^2\right)\left(1-\langle B_y\rangle^2\right)}}.
$$

1.9 Full-correlation approach to CHSH scenario

We now explore the CHSH scenario under the full-correlation perspective. All full-correlation boxes will be represented as column matrices:

$$
\boldsymbol{p}_F = \begin{bmatrix}\langle A_0B_0\rangle\\ \langle A_0B_1\rangle\\ \langle A_1B_0\rangle\\ \langle A_1B_1\rangle\end{bmatrix}.
$$

1.9.1 $\mathcal{FB}(2,2,2)$

The 16 vertices of the full-correlation general (non-signalling) CHSH scenario are just

$$
\boldsymbol{F}_{\langle A_0B_0\rangle\langle A_0B_1\rangle\langle A_1B_0\rangle\langle A_1B_1\rangle} = \begin{bmatrix}\langle A_0B_0\rangle\\ \langle A_0B_1\rangle\\ \langle A_1B_0\rangle\\ \langle A_1B_1\rangle\end{bmatrix}, \qquad (1.29)
$$

[40] In other words, these are *necessary*, but not *sufficient* conditions.

[41] Hierarchy that, in the asymptotic limit, provides inequalities that completely characterise the quantum box sets.

where $\langle A_x B_y \rangle \in \{-1, 1\}$. The $\mathcal{H}$-representation is given by the trivial inequalities $|\langle A_x B_y \rangle| \leq 1, \forall x, y$. We recall that all sets $\mathcal{FB}(N, I, 2)$ are hypercubes with edge size equals 2 and dimension I^N.

1.9.2 $\mathcal{FL}(2,2,2)$

We can easily find all the vertices of local full-correlation polytope by remembering that vertices can be reached by the full-correlation version of deterministic local boxes, so we necessarily have $\langle A_x B_y \rangle = \langle A_x \rangle \langle B_y \rangle$, with $\langle A_x \rangle \langle B_y \rangle \in \{-1, 1\}$. It is useful to define

$$\boldsymbol{L}_{\langle A_0 \rangle \langle A_1 \rangle \langle B_0 \rangle \langle B_1 \rangle} := \begin{bmatrix} \langle A_0 B_0 \rangle \\ \langle A_0 B_1 \rangle \\ \langle A_1 B_0 \rangle \\ \langle A_1 B_1 \rangle \end{bmatrix},$$

so we can obtain all local vertices by straightforward calculation:

$$\boldsymbol{L}_{++++} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{L}_{+++-} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{L}_{++-+} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{L}_{++--} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix};$$
$$\boldsymbol{L}_{+-++} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{L}_{+-+-} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{L}_{+--+} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{L}_{+---} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix};$$
$$\boldsymbol{L}_{-+++} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{L}_{-++-} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{L}_{-+-+} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{L}_{-+--} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix};$$
$$\boldsymbol{L}_{--++} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{L}_{--+-} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{L}_{---+} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{L}_{----} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Note that exactly the half of the vertices above are redundant. This happens because the full-correlation boxes are invariant over the substitution $A_x \mapsto -A_x$, $B_y \mapsto -B_y$ for all x and y. For example $\boldsymbol{L}_{----} = \boldsymbol{L}_{++++}$, $\boldsymbol{L}_{+--+} = \boldsymbol{L}_{-++-}$, *etc.*

The $\mathcal{H}$-representation of $\mathcal{FL}(2,2,2)$ is given by the trivial inequalities

$$-1 \leq \langle A_x B_y \rangle \leq 1,$$

and all 8 CHSH inequalities, which are already in the full-correlation form. We remark that for this specific scenario we have the property $\boldsymbol{p} \in \mathcal{L}(2,2,2) \iff FC(\boldsymbol{p}) \in \mathcal{FL}(2,2,2)$. This is not true in general, and in section 2.4 we will see that the CHSH scenario is the only multipartite scenario in which all Bell inequalities are in the full-correlation form.

It is also possible to check that $\mathcal{FL}(2,2,2)$ is a hyperoctahedron with edge size equals $\sqrt{2}$ in dimension 4 [51, 52, 53].

1.9.3 $\mathcal{FQ}(2,2,2)$

The extremal points and the inequality representation of $\mathcal{FQ}(2,2,2)$ were discovered independently by Landau (1988) [54], Tsirelson (1993) [13], and

Masanes (2003) [55]. The extremal points are

$$Q(\theta_1,\theta_2,\theta_3) = \begin{bmatrix} \sin(\theta_1) \\ \sin(\theta_2) \\ \sin(\theta_3) \\ -\sin(\theta_1+\theta_2+\theta_3) \end{bmatrix},$$

where $\theta_1, \theta_2, \theta_3 \in \mathbb{R}$, and the inequalities that completely characterise $\mathcal{FQ}(2,2,2)$ are

$$\begin{aligned}
-\pi &\leq -asin\langle A_0B_0\rangle + asin\langle A_0B_1\rangle + asin\langle A_1B_0\rangle + asin\langle A_1B_1\rangle \leq \pi; \\
-\pi &\leq +asin\langle A_0B_0\rangle - asin\langle A_0B_1\rangle + asin\langle A_1B_0\rangle + asin\langle A_1B_1\rangle \leq \pi; \\
-\pi &\leq +asin\langle A_0B_0\rangle + asin\langle A_0B_1\rangle - asin\langle A_1B_0\rangle + asin\langle A_1B_1\rangle \leq \pi; \\
-\pi &\leq +asin\langle A_0B_0\rangle + asin\langle A_0B_1\rangle + asin\langle A_1B_0\rangle - asin\langle A_1B_1\rangle \leq \pi.
\end{aligned} \tag{1.30}$$

We remark that this result, is in some sense, weaker than the CHSH inequalities. Of course, $\boldsymbol{p} \in \mathcal{Q}(2,2,2) \implies FC(\boldsymbol{p}) \in \mathcal{FQ}(2,2,2)$, but the converse is not true. There are non-quantum bipartite boxes that satisfy all 8 inequalities above [55, 13].

1.10 Geometrical aspects of box sets

We now summarize some geometrical properties on sets of multipartite boxes.

	dimension	vertices	facets
$\mathcal{B}(N,I,O)$	$(IO)^N - I^N$	$(O^N)^{(I^N)}$	$(IO)^N$
$\mathcal{NS}(N,I,O)$	$(I^O - I + 1)^N - 1$	?	$(IO)^N$
$\mathcal{Q}(N,I,O)$	$(I^O - I + 1)^N - 1$	∞	∞
$\mathcal{L}(N,I,O)$	$(I^O - I + 1)^N - 1$	$(IO)^N$	?

The dimension of these sets were calculated in [56, 19, 13]. The facets of $\mathcal{B}(N,I,O)$ and $\mathcal{NS}(N,I,O)$ are just the positivity inequalities, and the vertices of $\mathcal{L}(N,I,O)$ are the deterministic non-signalling boxes.

It is interesting to remark that the set $\mathcal{L}(N,2,2)$ is the polyhedral dual of the set $\mathcal{NS}(N,2,2)$, and this duality result only holds for $I = O = 2$ [57].

For full-correlations box sets we have

	dimension	vertices	facets
$\mathcal{FB}(N,I,2)$	I^N	2^{I^N}	$2I^N$
$\mathcal{FQ}(N,I,2)$	I^N	∞	∞
$\mathcal{FL}(N,I,2)$	I^N	?	?

Chapter 2

Developing some intuition on multipartite box sets

– Are you sure of that?
– Yes!... I think!
Roy and Maurice Moss, The IT crowd

This chapter is focused on presenting examples, results, and proof techniques for multipartite boxes. We hope that after reading this chapter, a non-specialist will gain some intuition on the structure of multipartite box sets.

2.1 Quantum violations of Bell inequalities

In this section we point out two necessary conditions on quantum systems for constructing nonlocal quantum boxes: quantum entanglement and quantum measurements that can be jointly performed. We start by defining quantum entanglement, that, as stated by one of the founding fathers of quantum mechanics[1], is one of its most particular features.

Since its extension to a multipartite scenario is simple, but with cumbersome notation, we will just present the definition for the bipartite case.

Definition 25 (Quantum entanglement [59])**.** *A quantum state $\rho_{AB} : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_B$ is separable if it can be written in the form*

$$\rho_{AB} = \sum_\lambda \pi(\lambda)\rho_A^\lambda \otimes \rho_B^\lambda$$

for some distribution $\pi : \Lambda \to [0,1]$ and quantum states $\rho_A^\lambda : \mathcal{H}_A \to \mathcal{H}_A$, $\rho_B^\lambda : \mathcal{H}_B \to \mathcal{H}_B$.

Quantum states that are not separable are entangled.

[1] Erwin Schrödinger [58]: "I would not call [quantum entanglement] one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought."

We now invite the readers to note that, as explicitly shown on theorem 6, the definition of separable states is very closely related to the definition of local boxes: quantum states are separable if they can be written as convex combination of product states.

In quantum information theory, entanglement is understood as a resource that can be used for protocols like, superdense coding [60], quantum teleportation [61], and possibly related to the exponential speed-up of quantum computation [62]. Now we will prove that quantum entanglement is also a necessary condition for constructing quantum nonlocal boxes.

Theorem 14. *Quantum boxes constructed by local measurements on separable states are local.*

Although this theorem holds for all multipartite scenarios, we will just present its proof for bipartite systems. We remark that the general proof follows from simple generalisation.

Proof. We just need to check that performing local measurements on separable states results in probabilities that can be described by a local boxes. Explicitly:

$$\begin{aligned} p_{ab|A_xB_y} &= \operatorname{tr} \rho_{AB} A_x^a \otimes B_y^b; \\ &= \sum_\lambda \pi(\lambda) \operatorname{tr} [(\rho_A^\lambda \otimes \rho_B^\lambda)(A_x^a \otimes B_y^b)]; \\ &= \sum_\lambda \pi(\lambda) \operatorname{tr} (\rho_A^\lambda A_x^a) \operatorname{tr} (\rho_B^\lambda B_y^b); \\ &= \sum_\lambda \pi(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda}. \end{aligned}$$

□

A natural question arises: "can all entangled states be used as a resource for constructing nonlocal boxes?". If the quantum state is a projector[2], the answer is positive [63], but we remark the existence of bipartite entangled states that cannot provide nonlocal correlations [64, 65] .

Another characteristic of quantum mechanics that is usually pointed as very particular is that many quantum measurements cannot be jointly performed. This characteristic is very closely related[3] to the fact that the measurement operators of two different quantum sets, $\{A^a\}, \{B^b\}$, may not obey the commutation relations $[A^a, B^b] = 0, \forall a, b$. In fact, it follows from a corollary of theorem 21 that if all quantum measurement sets of a given part obey the relation $[A_x^a, A_{x'}^{a'}] = 0$ with $x \neq x'$, its associated quantum box is local.

Theorem 15. *Let $\{A_x^a\}$ be a set of quantum measurements of one part of a bipartite quantum box. If the measurement operators respect $[A_x^a, A_{x'}^{a'}] = 0, \forall a, a'$ with $x \neq x'$ the associated quantum box is necessarily local.*

[2] In quantum mechanics, these states are known as pure states.

[3] For the cases in which measurement operators are projectors, two quantum measurement sets obey the commutation relation $[A^a, B^b] = 0, \forall a, b$ *iff* they can be jointly performed. But in general, the fact that two quantum measurements can be jointly performed does not imply that $[A^a, B^b] = 0, \forall a, b$ [66].

The idea behind this theorem is the fact that if the measurement operators of different sets commute, these measurements can be jointly performed. And if that happens, we have a single probability distribution that takes care of these outcomes simultaneously, so we can assign definite values for the outputs associated to this measurements.

We remark that this theorem also holds for the multipartite scenario: if the measurement operators of $(N-1)$ parties obey the commutation relation, the associated quantum box is local.

2.2 The relabelling transformation

One very useful tool in the study of multipartite boxes is the *relabelling transformation*, that was implicitly used in various papers but first formalised in [67]. A relabelling transformation on a set of box probabilities consists in permuting the label[4] of outputs and/or inputs. For example, in the scenario $(2,2,2)$ we have four different relabelling transformations:

- Alice performs an output relabelling: $p_{0b|A_xB_y} \leftrightarrow p_{1b|A_xB_y}$
- Alice performs an input relabelling: $p_{ab|A_0B_y} \leftrightarrow p_{ab|A_1B_y}$
- Bob performs an output relabelling: $p_{a0|A_xB_y} \leftrightarrow p_{a1|A_xB_y}$
- Bob performs an input relabelling: $p_{ab|A_xB_0} \leftrightarrow p_{ab|A_xB_1}$

The main motivation for this definition is the fact that a multipartite local (nonlocal) box cannot be transformed into a nonlocal (local) one just by the giving different names for the outputs and inputs.

This fact introduces some symmetries on the box sets $\mathcal{B}(\mathcal{S}), \mathcal{NS}(\mathcal{S}), \mathcal{Q}(\mathcal{S})$, and $\mathcal{L}(\mathcal{S})$.

We say that two equalities, inequalities, or boxes are in the same class if they are equivalent via relabelling transformation. For example, we can characterise all Bell inequalities of the $(2,2,2)$ scenario in two classes: trivial ones, that are equivalent to

$$p_{00|A_0B_0} \geq 0,$$

and the CHSH ones (see section 1.8), equivalent to[5]

$$\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle \leq 2.$$

Also, the set of all non-signalling vertices can be described by two classes: local ones, that are equivalent to

$$\boldsymbol{p}_L := \begin{bmatrix} p_{00|A_0B_0} & p_{00|A_0B_1} & p_{00|A_1B_0} & p_{00|A_1B_1} \\ p_{01|A_0B_0} & p_{01|A_0B_1} & p_{01|A_1B_0} & p_{01|A_1B_1} \\ p_{10|A_0B_0} & p_{10|A_0B_1} & p_{10|A_1B_0} & p_{10|A_1B_1} \\ p_{11|A_0B_0} & p_{11|A_0B_1} & p_{11|A_1B_0} & p_{11|A_1B_1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

[4] We also remark the existence of relabel between parties, that we do not push further due to the fact that they do not provide a different class of relabelling-equivalent boxes, inequalities, or equalities.

[5] To prove that, it may be useful to note that the transformation $p_{0b|A_xB_y} \leftrightarrow p_{1b|A_xB_y}$ implies in $\langle A_xB_y\rangle \mapsto -\langle A_xB_y\rangle$.

and nonlocal ones, equivalent to the PR-box

$$\boldsymbol{p}_{PR} := \begin{bmatrix} p_{00|A_0B_0} & p_{00|A_0B_1} & p_{00|A_1B_0} & p_{00|A_1B_1} \\ p_{01|A_0B_0} & p_{01|A_0B_1} & p_{01|A_1B_0} & p_{01|A_1B_1} \\ p_{10|A_0B_0} & p_{10|A_0B_1} & p_{10|A_1B_0} & p_{10|A_1B_1} \\ p_{11|A_0B_0} & p_{11|A_0B_1} & p_{11|A_1B_0} & p_{11|A_1B_1} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

2.3 Box purification

Deterministic boxes are the ones in which every outcome can be predicted with certainty. Such boxes admit a very simple description; we can write its outputs as a function of the inputs without mentioning probabilities. Motivated by the purification of quantum states [16], we will now try to understand an N-partite box as a deterministic $(N+1)$-partite one.

As an example, suppose Alice, Bob, and Eve share a deterministic box $\boldsymbol{p} \in \mathcal{B}(3,2,2)$, in which inputs and outputs[6] are defined by[7]:

$$\begin{aligned} o_A &= I_E, \\ o_B &= I_E + I_A I_B, \\ o_E &= I_A. \end{aligned} \tag{2.1}$$

If Eve presses her input I_E according to an uniform random distribution, Alice and Bob's statistics are exactly the ones obtained by a PR-box (more espeficly , $\boldsymbol{PR}_1$, see section 1.8.3), with an additional bonus that Eve learns Alice's input. Note as well that the outputs explicitly depend on others' input, fact that implies that this is a signalling box.

2.3.1 Understanding Alice's probabilities as Bob's manipulation

Suppose Alice has a non-deterministic box $\boldsymbol{p}^A \in \mathcal{B}(1, I_A, O_A)$. Can we understand the probabilistic nature of $\boldsymbol{p}^A$ as ignorance on Bob's input choices on a deterministic bipartite box? More formally, can we write

$$p_{a|A_x} = \sum_{b,y} \pi(y) p_{ab|A_xB_y} , \tag{2.2}$$

with $p_{ab|A_xB_y} \in \{0,1\}$, and $\pi : \mathcal{I}_B \to [0,1]$ representing the probabilities of Bob pressing B_y in a deterministic box $\boldsymbol{p}^{AB} \in \mathcal{B}(I_A, O_A; I_B, O_B)$?

Let us analyse the case in which Alice has a box with two inputs and two outputs, $\boldsymbol{p}^A \in \mathcal{B}(1,2,2)$, and we want to understand it as a marginal of[8] $\boldsymbol{p}^{AB} \in \mathcal{B}(2,2; I_B, 1)$ That is, we want to write

$$\boldsymbol{p}^A = \sum_y \pi(y) \begin{bmatrix} p_{0|A_0B_y} & p_{0|A_1B_y} \\ p_{1|A_0B_y} & p_{1|A_1B_y} \end{bmatrix} ,$$

[6] That take values on $\{0,1\}$.

[7] Here we are using addition modulo 2.

[8] We recall that $\boldsymbol{p}^{AB} \in \mathcal{B}(2,2; I_B, 1)$ is a box in which Alice has two inputs and two outputs per input, Bob has two inputs and one output per input. Please note that if one part has only one different output, we do not need to concern on the probability of obtaining it.

for some distribution on Bob's inputs $\pi : \mathcal{I}_B \to [0,1]$ and probabilities $p_{a|A_x B_y}$.

It may be clear that, if the number of Bob's inputs $\#\mathcal{I}_B = I_B$ is sufficiently large, the answer is yes. Just take I_B as the number of vertices of $\mathcal{B}(1,2,2)$ and write $\boldsymbol{p}_A$ as a convex combination of the vertices of $\mathcal{B}(1,2,2)$. In fact, Carathéodory theorem guarantees that if $I_B \geq d+1$, where d is the dimension of $\mathcal{B}(1,2,2)$, it is always possible to find one $\boldsymbol{p}^{AB}$ that purifies $\boldsymbol{p}^A$.

Please note that Bob's outputs are not important for proving the existence of this deterministic box, and Bob's input choices y plays the same role of the hidden variable λ (see section 1.14).

Does this result hold for all single part boxes? Can we always understand them as being marginals of deterministic bipartite boxes? Yes, but we warn that this deterministic box must[9] be signalling, if not, we would always have $p_{a|A_x;B_y} = p_{a|A_x}$ and

$$\sum_y \pi(y) p_{a|A_x} = p_{a|A_x}.$$

2.3.2 Multipartite purification

The result for box purification presented on the previous subsection also holds for the multipartite case, and its proof is completely analogous. For concreteness, before proving this fact, we will present some formal definitions.

Definition 26 (Purifying scenario)**.** *Let $\mathcal{S} = (N, I, O)$ be a multipartite scenario. We define an $(N+1)$-partite scenario where the $(N+1)$ part has I_P inputs and one output per input as*

$$\mathcal{S}_P := (\mathcal{S}; I_P). \tag{2.3}$$

We will refer to $\mathcal{S}_P$ as the purifying scenario associated to $\mathcal{S}$.

Definition 27 (Purifying box)**.** *Let $\mathcal{S} = (N, I, O)$ be an N-partite box scenario. An $(N+1)$-partite box $\boldsymbol{P} \in \mathcal{B}(\mathcal{S}; I_P)$ purifies a $\boldsymbol{p} \in \mathcal{B}(\mathcal{S})$ if we can write all probabilities of $\boldsymbol{p}$ as*

$$p_{\boldsymbol{o}|\boldsymbol{i}} := \sum_\lambda \pi(\lambda) P_{\boldsymbol{o}|\boldsymbol{i};\lambda} \tag{2.4}$$

with $P_{\boldsymbol{o}|\boldsymbol{i};\lambda} \in \{0,1\}$ for some distribution $\pi : \mathcal{I}_P \to [0,1]$.

Theorem 16 (All boxes can be purified)**.** *Let $\mathcal{S} = (N, I, O)$ be a multipartite box scenario and $\mathcal{S}_P = (\mathcal{S}; I_P)$ its associated purifying scenario. When $I_P \geq dim(\mathcal{B}(\mathcal{S})) + 1$, any N-partite boxes $\boldsymbol{p} \in \mathcal{B}(\mathcal{S})$ can be purified by an $(N+1)$-partite box $\boldsymbol{P} \in \mathcal{B}(\mathcal{S}_P)$.*

Proof. Since $\mathcal{B}(\mathcal{S})$ is a convex set, we can write all its elements by convex combination of its vertices (that are deterministic boxes). Then, we invoke Carathéodory's theorem (theorem 1) to ensure that we only need $I_P \geq \dim(\mathcal{B}(\mathcal{S})) + 1$. □

We remark that the dimension these box sets as function of N, I, and O are presented in section 1.10. And also, all purifying boxes $\boldsymbol{P}$ are signalling, or else $P_{\boldsymbol{o}|\boldsymbol{i};\lambda} = p_{\boldsymbol{o}|\boldsymbol{i}}$.

[9] Except for deterministic single boxes, which are already pure.

2.3.3 Box purification and nonlocality

Let us now illustrate one possible interpretation of box purification.

Imagine that Alice and Bob share a bipartite box $\boldsymbol{p} \in \mathcal{B}(2,2,2)$ that is manipulated by Eve's input choices λ. Explicitly,

$$\boldsymbol{p} = \sum_{\lambda} \pi(\lambda)\boldsymbol{p}^{\lambda},$$

where $\pi(\lambda)$ represents the probability that Eve chooses the input λ and the respective boxes

$$\boldsymbol{p}^{\lambda} := \begin{bmatrix} p_{00|A_0B_0;\lambda} & p_{00|A_0B_1;\lambda} & p_{00|A_1B_0;\lambda} & p_{00|A_1B_1;\lambda} \\ p_{01|A_0B_0;\lambda} & p_{01|A_0B_1;\lambda} & p_{01|A_1B_0;\lambda} & p_{01|A_1B_1;\lambda} \\ p_{10|A_0B_0;\lambda} & p_{10|A_0B_1;\lambda} & p_{10|A_1B_0;\lambda} & p_{10|A_1B_1;\lambda} \\ p_{11|A_0B_0;\lambda} & p_{11|A_0B_1;\lambda} & p_{11|A_1B_0;\lambda} & p_{11|A_1B_1;\lambda} \end{bmatrix},$$

which are the vertices[10] of $\mathcal{B}(2,2,2)$.

By theorem 16, all probabilistic boxes $\boldsymbol{p}$ can be understood as an ignorance on Eve's input choice. But what happens when $\boldsymbol{p}$ is nonlocal? If we recall that a box is local if and only if it can be written as a convex combination of deterministic non-signalling boxes, we are forced to admit that Eve *must* have access to signalling bipartite boxes to manipulate Alice and Bob's probabilities.

In some physical experiments[11], we have strong reasons to believe that signalling boxes cannot exist. The same reasoning suggests that probabilities cannot arise from a third party manipulation.

We remark that the above interpretation is intimately related to the *a priori* non-signalling loophole, discussed in section 4.1, and the security of *device independent protocols*, discussed in section 2.9.

2.4 More on the $(2,2,2)$ scenario

> The art of doing mathematics is finding that special case that contains all the germs of generality.
>
> David Hilbert

We now return to the $(2,2,2)$ scenario to analyse some of its particularities, and some characteristics that hold for all multipartite box scenarios. We warn that focusing too much on the CHSH scenario may lead us to some misguided intuitions, due to fact that some phenomena only take place in more complex scenarios, but since these scenarios are very complicated, it is useful to hold tight on a simple one and learn everything we can from it.

---

[10]We remark that the vertices of $\mathcal{B}(2,2,2)$ are deterministic boxes.

[11]Special relativity states that information takes a non-null time to propagate. So, if Alice and Bob are far apart, and the process of obtaining an output after pressing an input is very fast, there could be no signalisation.

2.4.1 The Fine theorem

In the same paper where Arthur Fine provided all non-trivial Bell inequalities for the CHSH scenario he also presented an alternative (but equivalent) definition for locality [42]. We now enunciate the theorem presented in Fine's original paper, and in section 3.4 we prove a generalisation of it, that provides an alternative definition for locality in all multipartite scenarios and also establishes a direct connection between locality and the *marginal problem* (see section 3.3).

Theorem 17 (Fine's theorem)**.** *A bipartite box in the CHSH scenario $\boldsymbol{p} \in \mathcal{B}(2,2,2)$ is local iff it is possible to construct a four-variable distribution with probabilities $P_{a_0a_1b_0b_1|A_0A_1B_0B_1}$ that can be used to recover all four distributions $p_{\cdot\,\cdot|A_xB_y}$ via the marginal calculations*

$$p_{a_0b_0|A_0B_0} = \sum_{a_1,b_1} P_{a_0a_1b_0b_1|A_0A_1B_0B_1}, \quad p_{a_0b_1|A_0B_1} = \sum_{a_1,b_0} P_{a_0a_1b_0b_1|A_0A_1B_0B_1};$$
$$p_{a_1b_0|A_1B_0} = \sum_{a_0,b_1} P_{a_0a_1b_0b_1|A_0A_1B_0B_1}, \quad p_{a_1b_1|A_1B_1} = \sum_{a_0,b_0} P_{a_0a_1b_0b_1|A_0A_1B_0B_1}.$$

As anticipated, an extension of this theorem holds for all multipartite scenarios: a box is local *iff* we can construct a “mother distribution” P that describes all inputs-outputs simultaneously. In other words, we can assign definite values to all outputs simultaneously *iff* the box is local.

2.4.2 CH inequalities

If we are concerned only with non-signalling boxes, we can use the normalisation and the non-signalling relations ((1.5) and (1.11)), to re-write the CHSH inequalities as

$$\begin{aligned}
-1 &\leq p_{00|A_0B_0} + p_{00|A_0B_1} + p_{00|A_1B_0} - p_{00|A_1B_1} - p_{0|A_0} - p_{0|B_0} \leq 0; \\
-1 &\leq p_{00|A_0B_1} + p_{00|A_0B_0} + p_{00|A_1B_1} - p_{00|A_1B_0} - p_{0|A_0} - p_{0|B_1} \leq 0; \\
-1 &\leq p_{00|A_1B_0} + p_{00|A_1B_1} + p_{00|A_0B_0} - p_{00|A_0B_1} - p_{0|A_1} - p_{0|B_0} \leq 0; \\
-1 &\leq p_{00|A_1B_1} + p_{00|A_1B_0} + p_{00|A_0B_1} - p_{00|A_0B_0} - p_{0|A_1} - p_{0|B_1} \leq 0, \qquad (2.5)
\end{aligned}$$

that, due to Clauser and Horn, are known as CH inequalities [68].

Note that the CH inequalities only have terms involving the output 0, a fact that can be very convenient in some situations[12]. Also, there is a useful theorem behind this representation. All non-signalling boxes can be completely represented by probabilities[13] involving $(O-1)$ outputs [19, 22]. For example, boxes with two outputs can always be represented by probabilities involving the output 0.

[12] Please see section 4.2.2 and theorem 25 for examples where the CH inequalities are more convenient than the CHSH ones.

[13] We warn that for using this representation, we also need to consider the marginals. For example, the CH inequalities involve the marginals $p_{a|A_x}$ and $p_{b|B_y}$.

2.4.3 CHSH in quantum mechanics

Recall that a bipartite box is quantum if we can write all probabilities as $p_{ab|A_xB_y} = \operatorname{tr}(\rho A_x^a \otimes B_y^b)$ (see definition 14). Since all Bell inequalities are in the form

$$\sum_{abxy} C(a,b,x,y) p_{ab|A_xB_y} \leq K,$$

we can explore the linearity of the trace function, to define a *Bell operator*

$$\boldsymbol{B} := \sum_{abxy} C(a,b,x,y) A_x^a \otimes B_y^b. \tag{2.6}$$

With a Bell operator $\boldsymbol{B}$, we can check if the quantum system with state ρ and measurement sets $\{A_x^a\}$, $\{B_y^b\}$ violates its associated Bell inequality simply by evaluating $\operatorname{tr}\rho\boldsymbol{B}$.

Bell operators are useful for analysing nonlocality on quantum systems, for example we can calculate the maximum quantum violation of a Bell inequality just by finding eigenvalues of a self-adjoint operator. We now illustrate some applications of Bell operators in the CHSH scenario.

Definition 28 (CHSH operators). *Let $\{A_i^0, A_i^1\}$ and $\{B_j^0, B_j^1\}$, $i,j \in \{0,1\}$, be measurement operator sets on different quantum systems. A CHSH operator is defined as*

$$CHSH := A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1,$$

where[14] *$A_i := A_i^0 - A_i^1$ and $B_j := B_j^0 - B_j^1$.*

We remark that the inequalities $-2 \leq \operatorname{tr}(\rho CHSH) \leq 2$ correspond to the CHSH inequalities (1.25d), and in quantum mechanics, the quantity

$$\langle A_x B_y \rangle_\rho := \operatorname{tr}(\rho A_x \otimes B_y)$$

is known as the expected value of operator $A_x \otimes B_y$ for the state ρ.

By diagonalising $CHSH$ operator we can find the maximum quantum violation of a Bell inequality and the states that attain it. And noting that CHSH inequalities are symmetric via (-1) multiplication, its maximal quantum violation is given by its norm[15]

$$\|CHSH\| := \max_{\rho \geq 0, \operatorname{tr}\rho = 1} \left| \langle CHSH \rangle_\rho \right|,$$

that can also be defined as the maximum singular value of the operator.

Theorem 18 (Tsirelson-Khalfin-Landau identity [47, 70]). *The above defined operators obey*

$$\|CHSH\|^2 \leq 4 + \|[A_0, A_1]\| \|[B_0, B_1]\|. \tag{2.7}$$

Moreover, if all measurement operators are projectors, equality holds.

[14] Please note that these are self-adjoint operators. In quantum mechanics self-adjoint operators are called *quantum observables*. A brief discussion on that is made on the beginning of chapter 4.

[15] Readers who are not familiar with norms of linear operators are invited to the section 2.4 7 of John Watrous online lecture notes [69], that discuss operator norms and some applications on quantum mechanics.

For the case that measurement operators are projectors, the proof of this theorem follows by straight forward calculation of $CHSH^2$. For the general case we can use an algebraic trick presented in [47].

Since $\|[A_0, A_1]\| \leq 2\|A_0 A_1\|$ and operators A_i have eigenvalues lying in $[-1, 1]$, we can obtain the Tsirelson bounds (inequalities (1.27)) with inequality (2.7).

Another important result that follows from inequality (2.7) is the fact that the maximum quantum violation can be attained with quantum systems lying in $\mathcal{H}_2 \otimes \mathcal{H}_2$, which is the simplest quantum system that can present nonlocality. This can be checked by setting $A_0 = B_0 = \sigma_z$, and $A_1 = B_1 = \sigma_x$, where σ_i is a Pauli matrix[16].

We also remark that with Tsirelson-Khalfin-Landau identity it is easy to check that if the measurement sets of one part obey the commutation relation $[A_0^a, A_1^{a'}] = 0$, the associated box is necessarily local[17].

We end this section by proving that if we "force" a certain quantum correlator to be maximum, we put some constraints on the maximal quantum violation.

Theorem 19 ([18]). *Let* $\boldsymbol{p} \in \mathcal{Q}(2,2,2)$ *be a quantum box in which the associated quantum state and all measurement operators are projectors. If* $|\langle A_i B_j \rangle| = 1$, *for any given* i, j, *then* $\max|\langle CHSH \rangle| = 5/2$.

In the proof of this theorem we will use some standard notation on quantum mechanics. Readers that are not used to it can find some explanation at the beginning of chapter 4.

Proof. Exploring the ideas presented in [72], we define

$$\begin{aligned} |A_0\rangle &:= A_0 \otimes \mathbb{1}|\psi\rangle, & |B_0\rangle &:= \mathbb{1} \otimes B_0|\psi\rangle; \\ |A_1\rangle &:= A_1 \otimes \mathbb{1}|\psi\rangle, & |B_1\rangle &:= \mathbb{1} \otimes B_1|\psi\rangle. \end{aligned}$$

So $\||A_i\rangle\| = \||B_i\rangle\| = 1$ and

$$\langle\psi|CHSH|\psi\rangle = \langle A_0|B_0\rangle + \langle A_0|B_1\rangle + \langle A_1|B_0\rangle - \langle A_1|B_1\rangle.$$

Now we choose $\langle A_0|B_0\rangle = 1$, the proof being the same for other i, j. So $|A_0\rangle = |B_0\rangle$ and we can write the expected value of the CHSH operator as

$$\begin{aligned} |\langle\psi|CHSH|\psi\rangle| &= |1 + \langle B_0|B_1\rangle + \langle A_1|B_0\rangle - \langle A_1|B_1\rangle| \\ &\leq |1 + \langle B_0|B_1\rangle| + |\langle A_1|(|B_0\rangle - |B_1\rangle)| \\ &\leq |1 + \langle B_0|B_1\rangle| + \||B_0\rangle - |B_1\rangle\| \\ &= |1 + \langle B_0|B_1\rangle| + \sqrt{2}\sqrt{1 - \langle B_0|B_1\rangle} \\ &\leq 5/2. \end{aligned}$$

[16]Pauli matrices are very useful in the study of linear operators of $\mathcal{H}_2$, and its definition can be found in the beginning of chapter 4.

[17]In fact, in [71] the authors prove that $|\mathrm{tr}\,\rho CHSH| \leq 2$ for all states ρ *iff* the measurements on one part can be jointly performed. This theorem absorbs the above statement on the commutation relation as a particular case.

[18]Proof obtained in collaboration with Mateus Araújo, also presented in the appendix of [1]

Note that $\langle B_0|B_1\rangle = \langle A_0|B_1\rangle$ is real, as an expected value of a self-adjoint operator, so we can pass from the third line to the fourth. □

We can generalize this theorem by fixing the value of $|\langle A_i \otimes B_j\rangle|$ and optimising with respect to the other correlation terms. By using this framework we can recover the above theorem, prove that if $|\langle A_i \otimes B_j\rangle| = 0$ and all quantum states and measurement operators are projectors, then $\max|\langle CHSH\rangle| = 3\sqrt{3}/2 \approx 2.60$, or prove that $|\langle A_i \otimes B_j\rangle| = 1/\sqrt{2}$ for all i, j is a necessary condition for attaining the Tsirelson bound. The general result is presented in figure 2.1.

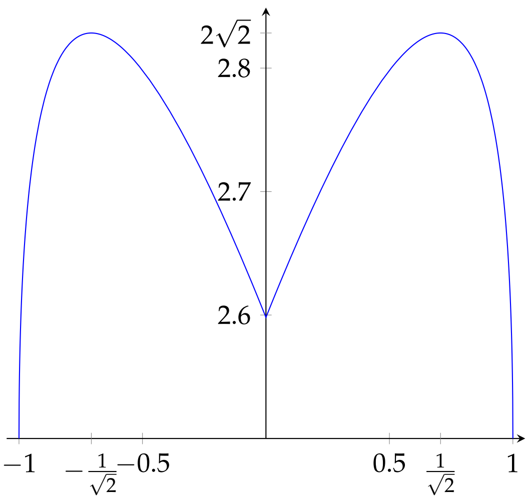

Figure 2.1: $\max|\langle CHSH\rangle|$ as a function of any fixed expected value.

2.5 The $(2, 3, 2)$ scenario

The bipartite scenario with three dichotomic measurements per party was first studied by Froissart in 1981 [37], who completely characterised its local polytope by tight Bell inequalities. We remark that Froissart's results were rediscovered by different authors in 2001, 2003, and 2004 [73, 74, 75] .

In this scenario, the facet representation of the local polytope has $4.(3.3) = 36$ positivity inequalities, $8\begin{pmatrix} 3 \\ 2 \end{pmatrix}^2 = 72$ CHSH inequalities, and 576 "new" inequalities relabelling-equivalent to[19]

$$I_{3322} := \langle A_1\rangle - \langle A_2\rangle + \langle B_1\rangle + \langle B_2\rangle + \langle A_1B_1\rangle + \langle A_2B_1\rangle + \langle A_3B_1\rangle \\ + \langle A_1B_2\rangle + \langle A_2B_2\rangle - \langle A_3B_2\rangle + \langle A_1B_3\rangle - \langle A_2B_3\rangle \leq 4. \quad (2.8)$$

The I_{3322} inequality has some properties that differ from the CHSH ones. First, it *cannot* be written in the full-correlation form, implying, that the local

[19]Please note that since $\langle X_i\rangle := p_{0|X_i} - p_{1|X_i}$, we are assuming non-signalisation to write I_{3322} in this simple form.

boxes of scenarios $(2, I, 2)$ with $I \geq 3$ cannot be completely characterised only by its full-correlators.

Also, differently from the CHSH inequalities, maximal quantum violation of I_{3322} inequalities cannot be attained by quantum systems lying in $\mathcal{H}_2 \otimes \mathcal{H}_2$, and the results contained in [21] suggest that maximal quantum violation can only be attained with infinite dimension vector spaces. This phenomenon will be discussed again on subsection 2.9.1.

Also, there are quantum states that do not violate any CHSH inequality, but violate I_{3322}. Indicating that the problem of classifying the quantum states that are useful for constructing nonlocal boxes may be very hard[20].

2.6 The $(3, 2, 2)$ scenario

The tripartite scenario with two dichotomic measurements per party was studied in [73, 74, 80].

In [73, 74] the authors characterise the local polytope by presenting 53856 inequalities organized in 46 different classes, and in [80] the authors analyse the non-signalling polytope by presenting its 53856 vertices split into 46 different classes. This "coincidence" is investigated by Tobias Fritz in [57], where he presents a simple bijection between the local and non-signalling polytopes. Moreover, he also show that the set $\mathcal{NS}(N, 2, 2)$ is the geometric dual of $\mathcal{L}(N, 2, 2)$.

We now point out some curious facts about this scenario.

- Some vertices of the non-signalling polytope do not violate maximally any Bell inequality [80].
- Some vertices of the non-signalling polytope violate maximally many different Bell inequalities [80].
- There is one class of non-trivial Bell inequality that cannot be violated by quantum boxes. These inequalities are related to the "Guess your neighbour's input" game [81, 82]
- One class of Bell inequality can be violated up to its algebraic maximum with quantum boxes[21], but no extremal point can be reached by performing measurements on quantum systems. These quantum boxes are related to the GHZ argument [83, 84].
- There are inequalities that cannot be written in the full-correlation form, implying that the local boxes of scenarios $(N, 2, 2)$ with $N \geq 3$ cannot be completely characterised only by its full-correlators.

[20] In [76], Asher Peres conjectured that a quantum state may be used for nonlocality *iff* it has a *negative partial transpose* [77]. For the multipartite case, this conjecture was proved false [78, 79], but the question remains open for the bipartite case.

[21] We recall that, due to Tsirelson bound (inequalities (1.27)), this is not possible for CHSH inequalities.

2.7 Some results on the (N, I, O) scenario

Although simple scenarios are important for our comprehension, we saw in last subsections that they may lead us to wrong conclusions. In order to put the readers closer to the nonlocality research field, we now list some results for more complicated scenarios.

- Using the results of last section, we conclude that the only non-trivial scenario[22] in which the local polytope can be completely characterised only by full-correlation inequalities is $(2, 2, 2)$.
- As we saw, in the $(3, 2, 2)$ scenario, there are non-trivial Bell inequalities that are not violated by quantum boxes. In [82], the authors show how to find this kind of inequalities, and also that they can only appear in scenarios with more than two parties. We also remark [85], where the authors relate these inequalities to the *local orthogonality* principle, that may be useful for studying quantum boxes.
- As in the $(3, 2, 2)$ scenario, there are examples of quantum boxes that violate non-trivial Bell inequalities up to its algebraic maximum, even with these quantum boxes *not* being extremal points of the non-signalling polytope. In [86], Adán Cabello shows that this phenomenon appear in the bipartite scenario with three inputs per party and four outputs per input $(2, 3, 4)$, and in [87], he establishes a direct connection among these boxes, *state independent contextuality* [88], and *quantum pseudo telepathy* [89].
- In [90] the authors provide a simple technique to generate non-trivial Bell inequalities for the $(2, 2, O)$ scenario. These inequalities are known as CGLMP inequalities, and they were proved to be tight in [67].
- In [91] the authors provide a simple technique to generate non-trivial Bell inequalities for the $(2, I, 2)$ scenario. These inequalities are known as Braunstein-Caves/chain inequalities, and they were proved to be *not* tight in [73, 74, 75].
- In [51, 52] the authors present all full-correlation Bell inequalities for the $(N, 2, 2)$ scenario. They also show that the set of full-correlation local boxes of this scenario is an hyper-octahedron, the geometric dual of the full-correlation non-signalling box set for the same scenario.

2.8 General properties of nonlocal boxes in non-signalling scenarios

In [92] the authors analyse general properties of nonlocal boxes in a framework where signalling boxes cannot exist. In special they prove some *monogamy conditions* that can be explored to guarantee security of some cryptographic

[22] A scenario is trivial if its local polytope is is equivalent to the non-signalling polytope. For examples on trivial scenarios, please see theorems 5 and 22.

protocols. For example, let $\boldsymbol{p} \in \mathcal{NS}(3,2,2)$ be a non-signalling tripartite box shared by Alice, Bob, and Eve. If the marginal box[23] of Alice and Bob is a nonlocal vertex of $\mathcal{NS}(2, I, O)$, the distributions between Alice and Eve are completely uncorrelated (see equation 1.6). And for the specific scenario where the three parties have two dichotomic inputs, $(3,2,2)$, if Alice and Bob marginal box is nonlocal them Alice and Eve marginal box is necessarily local.

We remark that in this same paper, the authors show that non-signalling physical theories that can be used for nonlocality have a *no-cloning theorem*, in the sense that the parties cannot make copies of an arbitrary state[24]. They also analyse the connection between nonlocality and other properties like intrinsic randomness and disturbance associated with measurements.

2.9 Device independent protocols

Since the beginning of this dissertation we have been studying properties of physical systems by considering only its statistical data. This approach can be used for derive *device independent* protocols, that are protocols in which its conclusions do not rely on the physical structure of the system but only on its statistics.

In 1991, Artur Ekert presented one cryptographic quantum protocol in which security was based on nonlocality arguments. The ideas presented on Ekert's paper were push further on [94, 95], where the authors show that in a scenario where signalling boxes cannot exist, nonlocality can be used for cryptographic protocols. Since then, other non-cryptographic protocols based on nonlocality were developed, for example, *random number certification* [96], *state estimation* [97], *entanglement measurement certification* [98], and *dimension witness*[25] [99].

We remark that in the physical world, the only way we know to construct nonlocal boxes is by performing measurements on quantum systems or by parties that can communicate with each other[26]. But we call the attention to the fact that, differently from other protocols based on quantum mechanics, the security of device independent ones hold independently of quantum theory. Fact that put the device independent protocols as one big motivator for physical realisation of nonlocal boxes, task that is showing to be more complicated than initially imagined [100]. For more on implementation of nonlocal boxes, please see chapter 4.

2.9.1 Dimension witness

We say that the dimension of a quantum system is the dimension of the smallest vector space sufficient for its description. An analogous definition

[23] A marginal box is the set of probabilities of the marginal distributions.

[24] We remark that the no-cloning theorem for quantum mechanics is known since 1982 [93], but before [92], its proof was based on the linearity of the vector space that describes quantum systems.

[25] Dimension witness will be discussed on next section

[26] We remark that nonlocal boxes that are constructed via communication between parties suffer from the *a priori* non-signalling loophole, and they cannot be used for device independent protocols. For more on that please see section 4.1 and the discussions on theorem 16.

exists for classical system, where the dimension is the number of variables of distribution necessary to define its physical states. The dimension of a physical system is sometimes associated to its complexity, and for quantum systems it plays a very important role on some cryptographic protocols.

As we saw in section 2.5, the amount of quantum violation of a given Bell inequality may depend on the dimension of its associated system. This phenomenon also happens for some inequalities in other scenarios, like the $(2,2,3)$, where the maximum quantum violation attainable for quantum systems lying in $\mathcal{H}_3 \otimes \mathcal{H}_3$ is strictly larger for quantum systems lying in $\mathcal{H}_2 \otimes \mathcal{H}_2$ [101].

In [99], the authors explore these results to develop the idea of *dimension witness*, an inequality that can be used to bound the minimum dimension of a physical (quantum or not) system just by analysing it statistical data. We remark that, in real physical experiments it may be hard to know exactly what is the minimum dimension necessary for describing its system.

More recently, the notion of dimension witness was explored[27], in a framework with two black boxes: one that can produce various different states, other that performs measurements [103]. We remark that in this new framework, the authors did not explore joint measurments, leaving a big room for improvement.

[27]Ideas that were already explored in a real physical laboratory [102].

Chapter 3

General box correlations

Sometimes, generalising makes the problem simpler.

Marcelo Terra Cunha

In chapter 1 we presented the definition of a multipartite box, which is a convenient way to represent I^N different N-variable distributions. We also discussed three cases of interest, non-signalling, quantum, and local boxes, and analysed some of their properties.

In this chapter we continue our studies on sets of multivariable distributions, but now in a more general framework, where we overcome the multipartite interpretation. We develop the concepts of *marginal condition* and *contextual boxes*, that play the same role of the non-signalling condition and nonlocal boxes. Also, in this general formalism, we will be able to discuss simple and natural scenarios that were missed behind the multipartite assumption.

We conclude the chapter presenting the results of [4], where, we[1] completely characterise the set of noncontextual boxes of the n-cycle scenario by presenting its noncontextuality inequalities. That is the first time that a family of noncontextual polytopes is described by its facets in a scenario with an arbitrary number of settings.

3.1 General boxes

Imagine that Alice has a black box in her laboratory which has 3 input buttons, labelled as I_1, I_2, and I_3. Each input has two associated outputs, 0 and 1. This box also respects one rule: it only works when Alice presses two buttons, that is, to receive an output Alice must choose two different inputs I_j and I_k.

The input/output probabilities of this box are given by three two-variable

[1]Mateus Araujo, Marco Túlio Quintino, Costantino Budroni, Marcelo Terra Cunha, and Adán Cabello.

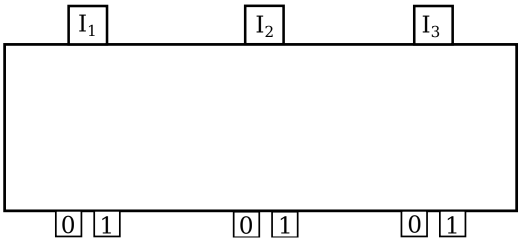

Figure 3.1: An illustration a 3-input dichotomic box scenario.

distributions, that can be represented as a vector $\boldsymbol{p} \in \mathbb{R}^{12}$,

$$\boldsymbol{p} = \begin{bmatrix} p_{00|I_1 I_2} & p_{00|I_2 I_3} & p_{00|I_3 I_1} \\ p_{01|I_1 I_2} & p_{01|I_2 I_3} & p_{01|I_3 I_1} \\ p_{10|I_1 I_2} & p_{10|I_2 I_3} & p_{10|I_3 I_1} \\ p_{11|I_1 I_2} & p_{11|I_2 I_3} & p_{11|I_3 I_1} \end{bmatrix}. \tag{3.1}$$

Please note that the above box scenario is very similar to the others presented on chapter 1, but with one important difference, we did not mention various different parties, the whole blackbox is under Alice's control. Note however that the non-signalling property defined for multipartite boxes has a clear analogue. If $p_{0|I_1;I_2} = p_{0|I_1;I_3}$, we can say that the inputs I_2 and I_3 do not "disturb"[2] I_1, and unambiguously define the marginal probability $p_{0|I_1}$. Locality also has its analogue, we can ask if it is possible to find a distribution $\pi : \Lambda \to [0, 1]$ and single variable distributions $p_{\cdot|I_{1;\lambda}}$, $p_{\cdot|I_{2;\lambda}}$, $p_{\cdot|I_{3;\lambda}}$ such that

$$p_{ab|I_x I_y} = \sum_{\lambda} \pi(\lambda) p_{a|I_{x;\lambda}} p_{b|I_{y;\lambda}}, \quad \forall a, b, x, y. \tag{3.2}$$

To motivate more the problem, we invite the readers to analyse the correlations of this box

$$\boldsymbol{p} = \begin{bmatrix} p_{00|I_1 I_2} & p_{00|I_2 I_3} & p_{00|I_3 I_1} \\ p_{01|I_1 I_2} & p_{01|I_2 I_3} & p_{01|I_3 I_1} \\ p_{10|I_1 I_2} & p_{10|I_2 I_3} & p_{10|I_3 I_1} \\ p_{11|I_1 I_2} & p_{11|I_2 I_3} & p_{11|I_3 I_1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.3}$$

Note that the outputs are always *anticorrelated*, that is, after pressing I_j and I_k, and checking that the button I_j outputted 0, she is sure that the output of I_k will be 1. Can probabilities of this be written on the form of of equation (3.2)? That is, can we understand its probabilistic nature a result from a unknown variable λ?

Well, let us assume that the outcomes of this box could be, in principle, deterministic. This means that we could assign a number, let us say 0, to the output of I_1. By the anticorrelation, we must assign 1 to the output I_2. Using the same argument, I_3 must be 0. But this same argument would imply

[2]We will see in section 4.1, that this condition may be harder to be *a priori* justified than the non-signalling condition. This fact may affect our interpretation of properties of general boxes.

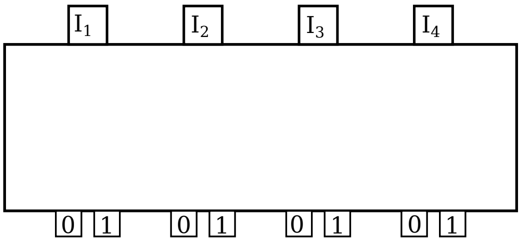

Figure 3.2: An illustration a 4-input dichotomic box scenario.

that I_1 must output 1, which is in contradiction with the assignment we first gave to I_1. This means that the box presented in equation (3.3) is *contextual*, a property that is analogous to nonlocality.

In order to establish a clear connection with the previous chapters, we now present one more example, the CHSH scenario (discussed in section 1.8) in this single party box formalism. Alice now has a black box with 4 inputs, labelled as I_1, I_2, I_3, and I_4. Each input can output 2 different values, 0 or 1. This box also respects some rules: to receive an output, Alice must press two "neighbour" inputs, that is $\{I_1, I_2\}$, $\{I_2, I_3\}$, $\{I_3, I_4\}$, or $\{I_4, I_1\}$. The probabilities that completely describe this box are

$$\boldsymbol{p} = \begin{bmatrix} p_{00|I_1I_2} & p_{00|I_2I_3} & p_{00|I_3I_4} & p_{00|I_4I_1} \\ p_{01|I_1I_2} & p_{01|I_2I_3} & p_{01|I_3I_4} & p_{01|I_4I_1} \\ p_{10|I_1I_2} & p_{10|I_2I_3} & p_{10|I_3I_4} & p_{10|I_4I_1} \\ p_{11|I_1I_2} & p_{11|I_2I_3} & p_{11|I_3I_4} & p_{11|I_4I_1}. \end{bmatrix}. \tag{3.4}$$

We now invite our readers to recognize that, except for different notation, the above matrix is exactly the same as (1.22) presented on section 1.8. And, as commented before, for different motivations, we can define properties that are analogous to non-signalling and locality.

To define a general box scenario, one needs to specify the number of inputs, the number of outputs and also a set of rules specifying which buttons can be pressed together. We remark that many authors refer to these rules as a context.

Definition 29 (General scenario)**.** *Let $\mathcal{I} = \{I_n\}_1^I$ be the set of inputs, $\mathcal{O} = \{o_n\}_1^O$ be the set of outputs per input and $\mathcal{R} = \{\mathcal{R}_j\}_1^R \subseteq 2^{\mathcal{I}}$, be the set of rules that define the inputs that can be pressed together. The 3-tuple $\mathcal{S} = (\mathcal{I}, \mathcal{O}, \mathcal{R})$ is called a general box scenario.*

Since the label of inputs and outputs have no deep meaning, we also use the shorthand notation $\mathcal{S} = (I, O, \mathcal{R})$ for the general box scenario with I inputs, O outputs per input and rules $\mathcal{R}$.

Definition 30 (General box)**.** *A general box in a scenario $(I, O, \{\mathcal{R}_j\}_1^R)$ is the set of R O-variable distributions that assign probabilities*

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} := p_{o_a o_b \ldots o_x | I_a I_b \ldots I_x}$$

to the outcomes o_a, o_b, $\ldots$, o_x when the inputs $I_a I_b \ldots I_x$ are pressed.

The set of all general boxes that can be created with I inputs, O outcomes per input, and rules $\mathcal{R}$ is $\mathcal{GB}(I, O, \mathcal{R})$.

It may be clear that all multipartite boxes are general boxes[3] but some general boxes cannot be understood as multipartite boxes[4]. So, this new formalism allows us to talk about interesting blackboxes that do not fit the multipartite assumption.

3.1.1 M-box

In section 1.3.1, the question "When can we talk about marginal probabilities without specifying other inputs?" motivated the definition of non-signalling boxes. As we have seen in the beginning of this chapter, the non-signalling condition has a clear analogue, but since this general framework does not rely on the different parties assumption, some authors prefer to refer to this property as *non-disturbance* [104], or *marginal condition*[5] [105, 106].

Definition 31 (M-box). *A general box is marginal if $\mathcal{R}$ is closed under the inclusion relation and all probabilities $p_{\mathbf{r}_j|\mathcal{R}_j}$ are consistent via marginals. That is $\mathcal{R}_k \in \mathcal{R}$ and $\mathcal{R}_j \subseteq \mathcal{R}_k$ implies $\mathcal{R}_j \in \mathcal{R}$, also*[6]

$$p_{\mathbf{r}_j|\mathcal{R}_j;\mathcal{R}_k\setminus\mathcal{R}_j} = \sum_{\mathcal{R}_k\setminus\mathcal{R}_j} p_{\mathbf{r}_k|\mathcal{R}_k} =: p_{\mathbf{r}_j|\mathcal{R}_j}. \tag{3.5}$$

The set of all marginal boxes that can be created with I inputs, O outcomes per input, and rules $\mathcal{R}$ is $\mathcal{M}(I, O, \mathcal{R})$.

The requirement that $\mathcal{R}$ is closed under inclusion is made just to guarantee that if Alice can press the inputs[7] $\{I_1, I_3, I_7\}$, she can also press these inputs individually ($\{I_1\}$, $\{I_3\}$, $\{I_7\}$), also two by two ($\{I_1, I_3\}$, $\{I_3, I_7\}$, $\{I_7, I_1\}$). The marginal/non-signalling/non-disturbance condition (equation (3.5)), ensures that when Alice presses the input $\{I_1, I_3\}$, the probability to obtain a given output cannot depend on $\{I_7\}$.

3.1.2 Noncontextual box

We will now define the local property for the general box scenario, but again, the world locality may not be appropriated in a single party scenario, so its analogue is usually referred to as *noncontextuality*. We call the attention to the fact that some older papers did not use the operational definition that will be presented. Our definition follows the same line of [56, 105, 106, 107, 87]. For some discussions on definitions of noncontextuality, we suggest [108].

---

[3] For example, all bipartite boxes of the set $\mathcal{B}(2, I, O)$ are general boxes that lie $\mathcal{GB}(I^2, O, \{\mathcal{R}_j\})$. We just need to set $I_i = A_{i-1}$, $I_{i+I} = B_{i-1}$, and rules $\mathcal{R}_j = \{A_k, B_l\}$.

[4] *e.g.* the box represented in equation (3.1).

[5] The reason for this second name will become apparent on section 3.3.

[6] The notation $\sum_{\mathcal{R}_k\setminus\mathcal{R}_j}$ stands for summing in all outputs corresponding to inputs that belongs to $\mathcal{R}_k$ but not in $\mathcal{R}_j$.

[7] More formally, this condition states that the existence of $p_{\ldots|I_1,I_3,I_7}$ implies that all other possible combinations involving these inputs also exists.

Definition 32 (Noncontextuality)**.** *A general box $\boldsymbol{p} \in \mathcal{GB}(I, O, \{\mathcal{R}_j\})$ is noncontextual if all probabilities can be written as*

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \sum_{\lambda} \pi(\lambda) \prod_{I_i \in \mathcal{R}_j} p_{o_i|I_i;\lambda}$$

for a certain distribution $\pi : \Lambda \to [0,1]$ and single variable distributions $p_{\cdot|I_i;\lambda}$.

The set of all noncontextual boxes that can be constructed with I inputs, O outcomes per input, and rules $\mathcal{R}$ is $\mathcal{NC}(I, O, \mathcal{R})$

Due to this definition, various authors like to state "Nonlocality is a particular case of contextuality", although this phrase can make sense, we remark the readers that a more precise statement is: "If a general box can be transformed into a multipartite box, nonlocality is *equivalent* to contextuality. If a general box cannot be transformed into a multipartite box, nonlocality is not well defined." We just remark that interpretations of physical realisations of contextual boxes may be different from the nonlocal ones. This happens because in a multipartite scenario, we may invoke some physical theories[8] to guarantee an *a priori* non-signalling condition. We will discuss this in detail in section 4.1.

3.1.3 Quantum general boxes

Like in the multipartite scenario (see sections 1.3.3 and 1.5.4), quantum general boxes are defined as boxes that can be constructed via measurements on quantum systems. This general framework has only a subtle difference, we do not have the tensor product structure any more, so we need a different approach to guarantee that two (or more) quantum measurements can be jointly performed.

Identifying quantum measurements that can be jointly performed is a fundamental problem in quantum theory, but for defining quantum boxes, we will just use the fact that if the elements of different quantum measurement sets commute pairwise, they can be jointly measured[9] [66].

Definition 33 (Quantum box)**.** *A general box $\boldsymbol{p} \in \mathcal{GB}(I, O, \{\mathcal{R}_j\})$ is quantum if all probabilities can be written as*

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \mathrm{tr}\,(\rho \prod_{\mathcal{I}_k \in \mathcal{R}_j} I_k^{o_k}) \qquad o_k \in \boldsymbol{r}_j,$$

for a certain quantum state $\rho : \mathcal{H} \to \mathcal{H}$ and quantum measurement sets $\{\mathcal{I}_i = \{I_i^o\}_{o=1}^{O}\}_{i=1}^{I}$ such that if $\mathcal{I}_i, \mathcal{I}_{i'} \in \mathcal{R}_j$, $i \neq i'$, then they respect the commutation relation $[I_i^o\ ,\ I_{i'}^{o'}] = 0$.

The set of all quantum boxes that can be constructed with I inputs, O outcomes per input, and rules $\mathcal{R}$ is $\mathcal{Q}(I, O, \mathcal{R})$.

---

[8]Some clever physical experiments use the fact that special relativity inputs constraints on the speed of propagation of information to guarantee that signalling boxes cannot exist.

[9]Readers that are interested in this problem are invited to reference [66], that analyses joint measurements on quantum systems in a clear and modern way.

One natural question is, does the definition of quantum general boxes coincide with the definition of quantum multipartite boxes in the cases where we can transform a general box into a multipartite box? For historical reasons, this question is known as *the Tsirelson problem*. For finite dimensional Hilbert spaces, the answer is yes, but a general answer is still not known[10] [109].

Before ending this section, we review some recent progress in the Tsirelson problem. In [113] it is proved that Tsirelson's problem is equivalent to a mathematical problem called *Connes' embebbing* and [107] establishes a direct connection with *Kirchberg's QWEP conjecture*[11]. Also, in [114] the authors propose a physical intuition on why the two definitions for quantum boxes should be equivalent.

3.1.4 The noncontextual polytope

The general box framework is a natural extension of the multipartite framework presented in chapter 1, and many proofs about general box sets are completely analogous to the ones on multipartite box sets. For example the general box sets also obey the hierarchic condition[12].:

$$\mathcal{NC}(\mathcal{S}) \subset \mathcal{Q}(\mathcal{S}) \subset \mathcal{M}(\mathcal{S}) \subset \mathcal{GB}(\mathcal{S}),$$

where $\mathcal{S} = (I, O, \mathcal{R})$ is a general box scenario Also, all these sets are convex, and except for the quantum set, they are polytopes. Since the proofs of these statements use techniques that are completely analogous to the ones used in the multipartite framework, we will not present them.

Since the set of noncontextual boxes, $\mathcal{NC}(I, O, \mathcal{R})$, is a polytope, it admits the halfspace representation, so we have *noncontextuality inequalities*.

Definition 34 (Noncontextuality inequality)**.** *The linear inequalities that are respected by all noncontextual boxes are called nocontextuality inequalities. If the*

[10] Curiously, Tsirelson claimed (without proving) that the definition for quantum multipartite boxes using the tensor product structure is equivalent than the one using commuting observables in 1993 [13]. But after being required for a formal proof in 2006, Tsirelson realized that he did not have one. In Tsirelson's own words [109]:

> My ideas about quantum Bell-type inequalities, published first in 1980, were scantily noted in 1980-1989 (only by L.Landau, S.Summers, R.Werner and A.Grib, and only the simplest case). Being discouraged, I published in 1993 a survey, without proofs, and quitted. One of the claims in that survey drew attention in 2006, and I was asked by A.Acin for the proof. To my crying shame, my would-be-proof failed badly. Trying to provide a kind of antidote against my toxic claim, I issued in 2006 the question, whether it is true or false, to the now discontinued Braunschweig website on open problems in quantum information theory (see archived copy [110], see also the first 29 problems in the arXiv [111]; my one was problem 33). I was not the first to ask this question, but probably the first to publish it [112], and now it is called "Tsirelson's problem" (rather than "Tsirelson's error").

[11] Actually, a positive answer to Kirchberg's Quantum Weak Expectation Property (QWEP) conjecture on tensor products of C*-algebras would imply a positive answer to this question for all bipartite scenarios. And an extended version of Tsirelson's problem is equivalent to the QWEP conjecture.

[12] In fact, for almost all general scenarios, we have the strict subset relation relation. As an example, for any scenario that is more elaborate than the ones presented in section 3.5, we have $\mathcal{NC}(\mathcal{S}) \subset \mathcal{Q}(\mathcal{S}) \subset \mathcal{M}(\mathcal{S}) \subset \mathcal{GB}(\mathcal{S})$.

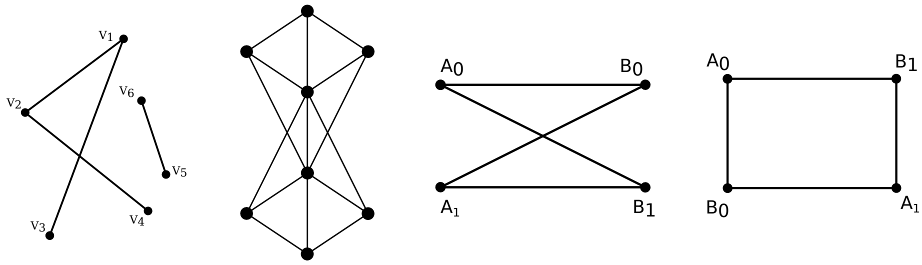

Figure 3.3: Some illustrations of graphs, with the two on the right representing the CHSH scenario (see section 1.8).

inequality is equivalent to the positivity, normalization, or non-signalling condition (equation (3.5)), we call it a trivial noncontextuality inequality. If a noncontextuality inequality represents a facet of a noncontextual polytope, it is called tight.

3.2 Full-correlation on general boxes

We will define the full-correlation map only for a special kind of general scenarios, the ones that can be represented by *graphs*. A graph $G = (V, E)$ is a set of vertices $V = \{V_i\}_{i=1}^{I}$ and a set of edges $E = \{V_i, V_j\}$, unordered pairs of vertices. For each graph, we can assign a general scenario by making the identification $\mathcal{I} = V$ and $\mathcal{R} = E$, and this scenario will only have pairwise rules. Since a graph does not mention the number of outputs per input, we adopt the convention that, if not specified, all inputs are dichotomic.

In section 1.7 we have defined the full-correlator map for bipartite dichotomic distribution $p_{\cdot\,\cdot|I_k I_k}$ as

$$\begin{aligned}\langle I_j I_k \rangle :&= p_{00|I_j I_k} + p_{11|I_j I_k} - p_{01|I_j I_k} - p_{10|I_j I_k} \\ &= p_{a=b|I_j I_k} - p_{a\neq b|I_j I_k}\end{aligned} \tag{3.6}$$

and by trivial extension, we defined the full-correlation map for bipartite boxes with two outcomes. Now, we present its version for general boxes that can be represented by graphs.

Definition 35. *Let $G = (V, E)$ be a graph and $\boldsymbol{p} \in \mathcal{GB}(V, 2, E)$ be a general box with two outputs and pairwise rules. The full-correlation box $FC(\boldsymbol{p})$ is the list of all correlators associated with $\boldsymbol{p}$.*

For example, in the scenario presented in equation (3.1), if

$$\boldsymbol{p} = \begin{bmatrix} p_{00|I_1 I_2} & p_{00|I_2 I_3} & p_{00|I_3 I_1} \\ p_{01|I_1 I_2} & p_{01|I_2 I_3} & p_{01|I_3 I_1} \\ p_{10|I_1 I_2} & p_{10|I_2 I_3} & p_{10|I_3 I_1} \\ p_{11|I_1 I_2} & p_{11|I_2 I_3} & p_{11|I_3 I_1} \end{bmatrix},$$

then

$$FC(\boldsymbol{p}) = \begin{bmatrix} \langle I_1 I_2 \rangle & \langle I_2 I_3 \rangle & \langle I_3 I_1 \rangle \end{bmatrix}.$$

Like in the multipartite framework, we can also define the full-correlation version of a given box set, and the set of all full-correlation general boxes is a hypercube with vertices' coordinates being ± 1.

3.3 The marginal problem

Instead of describing a box with a set of various distributions for specific rules, could we use a single distribution and recover all output probabilities as marginals? What could happen if Alice broke the rules and pressed all input buttons simultaneously? Is it possible to construct a distribution for this all input button situation which is consistent with the box probabilities?

Let us first focus on a specific scenario (see equation (3.1)). Given a box

$$\boldsymbol{p} = \begin{bmatrix} p_{00|I_1I_2} & p_{00|I_2I_3} & p_{00|I_3I_1} \\ p_{01|I_1I_2} & p_{01|I_2I_3} & p_{01|I_3I_1} \\ p_{10|I_1I_2} & p_{10|I_2I_3} & p_{10|I_3I_1} \\ p_{11|I_1I_2} & p_{11|I_2I_3} & p_{11|I_3I_1} \end{bmatrix},$$

when is it possible to construct a "mother" distribution

$$\boldsymbol{P} = \begin{bmatrix} P_{000|I_1I_2I_3} \\ P_{001|I_1I_2I_3} \\ P_{010|I_1I_2I_3} \\ P_{011|I_1I_2I_3} \\ P_{100|I_1I_2I_3} \\ P_{101|I_1I_2I_3} \\ P_{110|I_1I_2I_3} \\ p_{111|I_1I_2I_3} \end{bmatrix},$$

such that all probabilities of $\boldsymbol{p}$ can be obtained via the marginal calculations

$$\begin{aligned} p_{o_1o_2|I_1I_2} &= \sum_{o_3} P_{o_1o_2o_3|I_1I_2I_3}; \\ p_{o_2o_3|I_2I_3} &= \sum_{o_1} P_{o_1o_2o_3|I_1I_2I_3}; \\ p_{o_3o_1|I_3I_1} &= \sum_{o_2} P_{o_1o_2o_3|I_1I_2I_3}? \end{aligned}$$

More generally, given a box $\boldsymbol{p} \in \mathcal{G}(I, O, \mathcal{R})$, can we construct an I-variable distribution P in which all box probabilities can be obtained by calculating the marginal probabilities

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \sum_{\mathcal{I}\backslash\mathcal{R}_j} P_{\boldsymbol{o}|\mathcal{I}} \ ? \tag{3.7}$$

This problem of finding a mother distribution in which a set of distributions can be obtained as marginals is known (in mathematics) as the *marginal problem* and we will see in the next section that it is completely equivalent to the problem of deciding if a given box is noncontextual.

Although the connection of these two problems has been made, there is still a big gap between mathematicians and physicists. We illustrate this by quoting Alexander Klyachko (2002) [105]:

> The marginal problem has a long history, starting from works by W. Hoefling in Germany (1940), and a bit later by Freché in France. Springer Verlag published collected papers of Hoefling in 1994. Three conferences on the subject were held in the last decade [115, 116, 117]. None of the participants ever mentioned Bell's problem, and apparently none of the physicists was aware about these activities. This is a disturbing example of a split between mathematics and physics.

3.4 The Fine theorem

In 1982, Arthur Fine established a direct connection between local boxes in the CHSH scenario and the marginal problem [42]. Fine was only concerned with the CHSH scenario, but his theorem can be extended to all general box scenarios with minor modifications. We will now present a generalization of Fine's theorem, which can be also found in [56].

Theorem 20 (Arthur Fine)**.** *Let $\boldsymbol{p} \in \mathcal{M}(I, O, \mathcal{R})$ be a box with inputs $\mathcal{I} = \{\mathcal{I}_n\}_{j=1}^{I}$, O outputs per input and rules $\mathcal{R} = \{\mathcal{R}_j\}_{i=1}^{R}$. The box probabilities admit the representation*

$$p_{\boldsymbol{r}|\mathcal{R}_j} = \sum_{\lambda} \pi(\lambda) \prod_{I_i \in \mathcal{R}_j} p_{o_i|I_i;\lambda} \tag{3.8}$$

for some distribution $\pi : \Lambda \to [0, 1]$ and distributions $p_{.|I_i,\lambda}$ iff there exists a probability distribution $P_{.|\mathcal{I}}$ that can recover all box probabilities as marginals. That is

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \sum_{\mathcal{I} \setminus \mathcal{R}_j} P_{\boldsymbol{o}|\mathcal{I}}. \tag{3.9}$$

Proof. The theorem will be proved in a constructive way, we will construct $P_{.|\mathcal{I}}$ when π and $p_{o_i|I_i;\lambda}$ are given, and construct all $p_{\boldsymbol{s}|\mathcal{S}}$ and π when $P_{\mathcal{I}}$ is given.

If

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \sum_{\lambda} \pi(\lambda) \prod_{I_i \in \mathcal{R}_j} p_{o_i|I_i;\lambda},$$

we define

$$P_{\boldsymbol{o}|\mathcal{I}} := \sum_{\lambda} \pi(\lambda) \prod_{i} p_{o_i|I_i;\lambda},$$

and check that

$$\begin{aligned}\sum_{\mathcal{I}\setminus\mathcal{R}_j} P_{\boldsymbol{o}|\mathcal{I}} &= \sum_{\mathcal{I}\setminus\mathcal{R}_j} P_{o_1 o_2 \ldots o_I | I_1 I_2 \ldots I_I};\\ &= \sum_{\lambda}\sum_{\boldsymbol{o}\setminus\boldsymbol{o}_j} \pi(\lambda)\prod_i p_{o_i|I_i;\lambda};\\ &= \sum_{\lambda} \pi(\lambda)\prod_{\mathcal{R}_j} p_{o_i|I_i;\lambda};\\ &= p_{\boldsymbol{r}_j|\mathcal{R}_j}.\end{aligned}$$

If we have the distribution $P : \mathcal{I} \to [0,1]$, we invoke the multipartite generalisation of theorem 3 to write P as

$$P_{\boldsymbol{o}|\mathcal{I}} = \sum_{\lambda} \pi(\lambda)\prod_i p_{o_i|I_i\lambda},$$

and define the marginals as

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} := \sum_{\lambda} \pi(\lambda)\prod_{\mathcal{R}_j} p_{o_i|I_i;\lambda},$$

and check that

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \sum_{\mathcal{I}\setminus\mathcal{R}_j} P_{\boldsymbol{o}|\mathcal{I}}.$$

□

So, Fine's theorem provides an alternative definition/interpretation for contextuality. Informally it allows us to state that a box is noncontextual *iff* all possible outcome probabilities can be consistently described by a single distribution that attributes probabilities to all inputs simultaneously.

We will now use Fine theorem as tool to prove two important theorems.

Theorem 21. *Let $\mathcal{S} = (I, O, \mathcal{R})$ be a general scenario, and $\boldsymbol{p} \in \mathcal{Q}(\mathcal{S})$ be a quantum box in this scenario. If* all *measurement operators I_i^o satisfy the commutation relation $[I_i^o, I_{i'}^{o'}] = 0$, $\forall o, o'$ and $i \neq i'$, the box $\boldsymbol{p}$ is necessarily noncontextual.*

Proof. For any quantum state ρ we can define the function

$$P_{o_{i_1},\ldots,o_{i_I}|i_1\ldots i_I} := \operatorname{tr}\rho\prod_i I_i^{o_i}.$$

The normalisation relations on the measurement operators

$$\sum_{o_i} I_i^{o_i} = I,$$

guarantees the normalisation relation on the distribution

$$\sum_{\mathcal{O}} P_{o_{i_1},\ldots,o_{i_I}|i_1\ldots i_I} = 1,$$

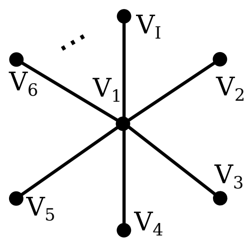

Figure 3.4: An illustration of a star graph.

and the commutative relation between measurement operators guarantees the positiveness[13] of P.

Now we just need to use the normalisation relation on the measurement operators to check that we can recover all box probabilities via the marginal calculation

$$p_{\boldsymbol{r}_j|\mathcal{R}_j} = \sum_{\mathcal{I}\setminus\mathcal{R}_j} P_{\boldsymbol{o}|\mathcal{I}}.$$

So, using Fine's theorem, we see that the box $\boldsymbol{p}$ is noncontextual.

□

We will now prove that all marginal boxes on the *star graph* scenario are noncontextual. A *star graph* is any graph that can be written as $G_S = (\{V_i\}, \{V_1, V_j\})$.

Theorem 22. *All marginal boxes with rules given by $\{I_1, I_j\}_{j=2}^{I}$ are noncontextual.*

Proof. Since $\boldsymbol{p}$ is a marginal box, the marginal $p_{o_1|I_1}$ is well defined, which allows us to construct

$$P := \frac{p_{o_1o_2|I_1I_2}\ p_{o_1o_3|I_1I_3}\ p_{o_1o_4|I_1I_4}\ \cdots\ p_{o_1o_I|I_1I_I}}{(p_{o_1|I_1})^{I-1}}.$$

Now, it is straightforward to check that P is a valid distribution and that we can recover the marginals via

$$p_{o_1o_k|I_0I_k} = \sum_{o_j\neq o_k} \frac{p_{o_1o_2|I_1I_2}\ p_{o_1o_3|I_1I_3}\ p_{o_1o_4|I_1I_4}\ \cdots\ p_{o_1o_I|I_1I_I}}{(p_{o_1|I_1})^{I-1}}.$$

□

Also note that this result also proves theorem 5. This presented theorem on star graphs can also be viewed as a particular case of a more general one. In [118], it is shown that if a general scenario can be associated to s graph without closed loops, the noncontextuality conditions are equivalent to the marginal conditions. This motivates us to explore cyclic graphs.

[13] If $A, B \geq 0$ and $[A, B] = 0$, $AB \geq 0$.

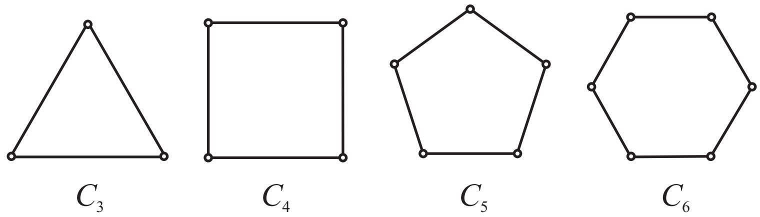

Figure 3.5: Some illustrations of cyclic graphs C_n.

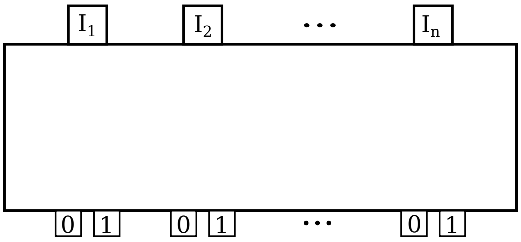

Figure 3.6: An illustration of the n-cycle box.

3.5 All noncontextuality inequalities for the two outcome n-cycle

The examples given in equation (3.1) and equation (3.4) can be generalised to a simple scenario called the n-cycle. A box belongs to the n-cycle if it has n inputs that outcome an output only when two neighbouring buttons are pressed.

The n-cycle scenario can also be represented by the *cyclic graph*[14], $C_n = (\{I_j\}_{j=1}^n, \{I_j, I_{j+1}\}_{j=1}^n)$. We define the dichotomic n-cycle scenario scenario as $\mathcal{C}_n := (n, 2, C_n)$, and its general boxes can be represented in a vector $\boldsymbol{p} \in \mathbb{R}^{4n}$,

$$\boldsymbol{p} = \begin{bmatrix} p_{00|I_1 I_2} & p_{00|I_2 I_3} & \cdots & p_{00|I_j I_{j+1}} & \cdots & p_{00|I_n I_1} \\ p_{01|I_1 I_2} & p_{01|I_2 I_3} & \cdots & p_{01|I_j I_{j+1}} & \cdots & p_{01|I_n I_1} \\ p_{10|I_1 I_2} & p_{10|I_2 I_3} & \cdots & p_{10|I_j I_{j+1}} & \cdots & p_{10|I_n I_1} \\ p_{11|I_1 I_2} & p_{11|I_2 I_3} & \cdots & p_{11|I_j I_{j+1}} & \cdots & p_{11|I_n I_1}. \end{bmatrix}.$$

In this section we present all noncontextuality inequalities for the dichotomic[15] n-cycle scenario. This result was obtained in collaboration with Mateus Araújo, Marcelo Terra Cunha, Costantino Budroni and Adán Cabello. We warn the readers that although all results presented in this section are contained in reference [4], here we reach the results from a different approach, using different proofs and lemmas.

[14] Here, the symbol + stands for addition modulo n, so $\{I_n, I_{n+1}\} = \{I_n, I_1\}$.
[15] And from now on, we will refer to the dichotomic n-cycle just as n-cycle.

3.5.1 Previous efforts on the n-cycle

The n-cycle scenario has been investigated long before physicists developed an interest in it. For instance, in 1862 Boole characterised the case $n = 2$ and studied the case $n = 3$ [29, 39], that was not discussed until almost a hundred years later [119, 120] and was only completely characterised in 1989 [22].

As stated before, the 4-cycle is completely equivalent to the CHSH scenario, which was characterised by Fine in 1982. The 5-cycle was explored by Klyachko in 2002 [105] and completely characterised by Klyachko *et al* in 2008 [121]. Klyachko's papers are also the first ones that use noncontextuality inequalities in the language that is used today. They also became very famous due to the fact that it is possible to construct 5-cycle contextual boxes using quantum systems with 3 dimensional Hilbert space[16].

More recently, there has been renewed interest in this scenario: noncontextuality inequalities have been found for any odd n [120, 122], entropic inequalities (which are necessary but not sufficient for noncontextuality) for any n [107, 106], and noncontextuality inequalities for $n = 6$ [87].

3.5.2 Main result

The main result of this section is theorem 23, which presents all tight noncontextuality inequalities for the n-cycle. And due to its simple form, it is very easy to tell if an n-cycle box is contextual or not.

Theorem 23 ($\mathcal{M}(\mathcal{C}_n)$ inequality characterisation)**.** *All tight non-trivial noncontextuality inequalities for the n-cycle are given by*

$$\sum_{j=1}^{n} \gamma_j \langle I_j I_{j+1} \rangle \leq n - 2, \tag{3.10}$$

where $\gamma_j \in \{-1, 1\}$, such that the number of negative coefficients γ_j is odd.

Please note that, like in the CHSH scenario (see 1.8 and 1.9), all inequalities are written in the full-correlation form (see equation (3.6)). In the multipartite framework, the scenario with 2 parties and 2 dichotomic inputs per party is the only one in which this is possible (see section 2.4.). This full-correlation symmetry on the n-cycle is probably what allows us to find such a simple characterisation. We recall that the problem of completely characterising the set of noncontextuality inequalities for an arbitrary number of settings is, in general, intractable, since deciding whether or not a given point is in the polytope is an NP-complete problem [22].

It is also interesting to remark that until now, there had been no complete characterisations of a local/noncontexutal polytope of a scenario with an arbitrary number of settings. The best result we know in this sense are the WWZB inequalities [51, 52] (see section 2.7), that characterise the *full-correlation* set of the multipartite scenario with n parties, and two dichotomic inputs per part.

[16] Before that, the smallest Hilbert space dimension required was 4.

3.5.3 Proofs

In order to present a clear demonstration and to gain some intuition on why this scenario is so simple, we divide the proof of theorem 23 in to eight lemmas.

- Lemma 1 proves two useful representations for boxes that lie in the marginal polytope of dichotomic scenarios.
- Lemma 2 characterises the contextual and noncontextual vertices of the full-correlation polytope for the n-cycle scenario.
- Lemmas 3 and 4 characterise the contextual and noncontextual vertices of the marginal polytope for the n-cycle scenario.
- Lemma 5 proves that all inequalities (3.10) represent facets of the noncontextual polytope of the n-cycle scenario.
- Lemma 6 shows that two contextual vertices of the marginal polytope of the n-cycle cannot be connected by edges of the marginal polytope.
- Lemma 7 presents a general characterisation for convex polytopes that uses disjoint union of a "core" polytope and some quasi-polytopes.
- Lemma 8 uses lemmas 6 and 7 to show that all contextual boxes of the marginal polytope of n-cycle scenario violate one of inequalities (3.10).

So, our main result follows directly from lemmas 5 and 8.

Lemma 1 (CH representations[17]). *Boxes that lie on the marginal polytope of a dichotomic scenario with pairwise rules, $\mathcal{M}(I, 2, \{I_j, I_k\}_{jk})$, are completely described by probabilities $p_{00|I_jI_k}$ and $p_{0|I_j}$, or by its full and single correlators $\langle I_jI_k\rangle$ and $\langle I_j\rangle$.*

Proof. To prove the CH representation for dichotomic scenarios with pairwise rules we just need to use the normalisation and non-signalling conditions to write all probabilities as functions of $p_{00|I_jI_k}$ and $p_{0|I_j}$. More explicitly,

$$p_{01|I_jI_k} = +p_{0|I_j} - p_{00|I_jI_k}; \tag{3.11a}$$

$$p_{10|I_jI_k} = +p_{0|I_k} - p_{00|I_jI_k}; \tag{3.11b}$$

$$\begin{aligned} p_{11|I_jI_k} &= 1 - p_{00|I_jI_k} - p_{01|I_jI_k} - p_{10|I_jI_k}, \\ &= 1 + p_{00|I_jI_k} - p_{0|I_j} - p_{0|I_k}. \end{aligned} \tag{3.11c}$$

To prove the correlator representation, we just need to show that it is possible to obtain probabilities $p_{00|I_jI_k}$ and $p_{0|I_j}$ with $\langle I_jI_k\rangle$ and $\langle I_j\rangle$. By definition $\langle I_j\rangle := p_{0|I_j} - p_{1|I_j}$, so using the normalisation condition, we have

$$p_{0|I_j} = \frac{\langle I_j\rangle + 1}{2}. \tag{3.12a}$$

[17] The name CH representation is inspired by the CH inequalities. For more, please see section 2.4.2.

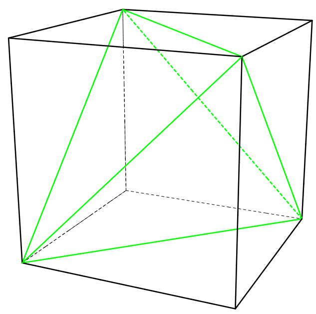

Figure 3.7: One tetrahedron and one cube representing the noncontextual and the general full-correlation 3-cycle polytope.

The definition of the full-correlator, $\langle I_j I_k \rangle := p_{00|I_j I_k} + p_{11|I_j I_k} - p_{01|I_j I_k} - p_{10|I_j I_k}$, together with the normalisation condition can be used to write $1 + \langle I_j I_k \rangle = 2p_{00|I_j I_k} + 2p_{11|I_j I_k}$. With equations (3.11c) and (3.12a) we obtain

$$p_{00|I_j I_k} = \frac{1 + \langle I_j I_k \rangle + \langle I_j \rangle + \langle I_k \rangle}{4}. \tag{3.12b}$$

□

Lemma 2 ($FC(\mathcal{M}(\mathcal{C}_n))$ vertex characterisation)**.** *The full-correlation n-cycle polytope is an n-dimensional hypercube with vertices with correlators $\langle I_j I_{j+1} \rangle = \pm 1$.*

Moreover, there are 2^{n-1} noncontextual vertices, all with an even *number of negative correlators, and relabelling-equivalent to*

$$\boldsymbol{NC}_F = \left[\langle I_1 I_2 \rangle = 1 \quad \langle I_2 I_3 \rangle = 1 \quad \langle I_3 I_4 \rangle = 1 \quad \ldots \quad \langle I_N I_1 \rangle = 1 \right].$$

There are 2^{n-1} contextual vertices, all with an odd *number of negative correlators, and relabelling-equivalent to*

$$\boldsymbol{C}_F = \left[\langle I_1 I_2 \rangle = -1 \quad \langle I_2 I_3 \rangle = 1 \quad \langle I_3 I_4 \rangle = 1 \quad \ldots \quad \langle I_N I_1 \rangle = 1 \right].$$

Proof. The full-correlation polytope of the n-cycle is an n-dimensional hypercube because all correlators $\langle I_j I_{j+1} \rangle$ can assume the values ± 1.

Before characterising the vertices as contextual and noncontextual, we recall that the output relabelling transformation[18], $0 \leftrightarrow 1$, cannot transform a noncontextual vertex into a contextual one[19]. We also point that the output

[18]We recall that for studying full-correlation boxes, it is useful to use the outputs $+1$ and -1 instead of 0 and 1. So we can understand the correlators as expected values.

[19]If it was possible, Alice could transform a noncontextual box into a contextual one just by giving different names to her outputs.

relabelling on the input I_j will be conveniently referred as $I_j \mapsto -I_j$, since this transformation changes the sign of the two "neighbour" correlators,

$$I_j \mapsto -I_j \implies \langle I_j I_{j+1} \rangle \mapsto -\langle I_j I_{j+1} \rangle \text{ and } \langle I_{j-1} I_j \rangle \mapsto -\langle I_{j-1} I_j \rangle.$$

We now classify the vertices in two classes, the ones that have an *even* number of correlators equal to -1, relabelling-equivalent to the vertex in which all correlators are equal to $+1$. The ones that have an *odd* number of correlator equal to -1, relabelling-equivalent to the vertex in which all correlators, except from $\langle I_1 I_2 \rangle = -1$, are equal to $+1$. We will present a simple algorithm to prove this equivalence.

First, note that the relabelling transformation allows us to "move" minus signs, and to cancel a pair of negative correlators. More precisely:

- If two neighbour correlators have negative values, $\langle I_{j-1} I_j \rangle = \langle I_j I_{j+1} \rangle = -1$, the map $I_j \mapsto -I_j$ transforms both into positive $\langle I_{j-1} I_j \rangle = \langle I_j I_{j+1} \rangle = +1$.
- If two neighbour correlators have different correlators, $\langle I_{j-1} I_j \rangle = +1, \langle I_j I_{j+1} \rangle = -1$, the map $I_j \mapsto -I_j$ "moves" the minus correlator $\langle I_{j-1} I_j \rangle = -1, \langle I_j I_{j+1} \rangle = +1$.

Now, let

$$\boldsymbol{V}_F = \left[\langle I_1 I_2 \rangle \quad \langle I_2 I_3 \rangle \quad \ldots \quad \langle I_{N-1} I_N \rangle \quad \langle I_N I_1 \rangle \right],$$

be a vertex of the full-correlation n-cycle polytope, for $k = N$ until $k = 2$, apply the relabelling transformation $I_j \mapsto -I_j$ when $\langle I_k I_{k+1} \rangle = -1$. This algorithm works because:

- If $\langle I_k I_{k+1} \rangle = -1$ and its left correlator neighbour is negative, both will become positive;
- If $\langle I_k I_{k+1} \rangle = -1$ and its left correlator neighbour is positive, the minus sign will move;

so this algorithm transforms any vertex into one that, except from $\langle I_1 I_2 \rangle$, has only positive correlators, dividing them in the two relabelling-equivalent classes discussed above.

We now recall that a vertex is noncontextual *iff* we can understand its inputs I_j as deterministic probability distributions. So, for noncontextual vertices, we can write $\langle I_j I_{j+1} \rangle = \langle I_j \rangle \langle I_{j+1} \rangle$ assigning definite values to each input. So, it follows that the vertex in which all its correlators $\langle I_j I_{j+1} \rangle$ are equal to $+1$ is noncontextual because we can assign the value 1 to every input, and $\langle I_j \rangle = 1, \forall j$.

Now consider the vertex which has, except from $\langle I_1 I_2 \rangle = -1$, all correlators equal to $+1$. If we assume that this vertex is noncontextual, $\langle I_1 I_2 \rangle = -1$ implies $(\langle I_1 \rangle = 1, \langle I_2 \rangle = -1)$ or $(\langle I_1 \rangle = -1, \langle I_2 \rangle = 1)$. Without loss of generality, we assume that $(\langle I_1 \rangle = 1, \langle I_2 \rangle = -1)$, what forces us to set

$\langle I_3 \rangle = -1$, because $\langle I_2 I_3 \rangle = 1$. By using this argument recursively, we find that we need $\langle I_1 \rangle = -1$, an absurdity.

□

Lemma 3 (Injection between noncontextual vertices)**.** *Let $\boldsymbol{p}_F = FC(\boldsymbol{p})$ be a noncontextual vertex of the full-correlation n-cycle polytope. There are exactly two noncontextual vertices of the marginal polytope such that its full-correlator part is equal to $\boldsymbol{p}_F$.*

Moreover, all noncontextual vertices of the marginal polytope can be described by a full-correlator noncontextual vertex and one single correlator $\langle I_k \rangle$. Also, if one noncontextual vertex $\boldsymbol{p}^a$ satisfies $\boldsymbol{p}_F = FC(\boldsymbol{p}^a)$, and $\langle I_k \rangle_{\boldsymbol{p}^a} = \pm 1$, the other associated vertex $\boldsymbol{p}^b$ that satisfies $\boldsymbol{p}_F = FC(\boldsymbol{p}^b)$ has $\langle I_k \rangle_{\boldsymbol{p}^b} = \mp 1$

Proof. Since all noncontextual vertices of the full-correlation polytope are relabelling-equivalent to the one in which all its full-correlators are equal $+1$ (lemma 2), we will just show that this specific vertex corresponds to exactly two noncontextual vertices of the marginal polytope.

We now recall that a box is a vertex of the noncontextual polytope *iff* we can assign definite values to every input. So, there are only two vertices that are consistent with all full-correlators being positive: the ones in which all inputs are 0, and the ones in which all inputs are 1. Also invoking lemma 1, we can represent one vertex as[20]

$$\langle I_j I_{j+1} \rangle = 1, \forall j \quad \text{and} \quad \langle I_1 \rangle = 1$$

and the other by

$$\langle I_j I_{j+1} \rangle = 1, \forall j \quad \text{and} \quad \langle I_1 \rangle = -1.$$

Since there are 2^n ways to assign definite outputs to all inputs simultaneously, we can use lemma two to guarantee that these are the only noncontextual vertices of the marginal polytope.

□

Lemma 4 (Bijection between contextual vertices)**.** *A full-correlation box $\boldsymbol{p}_F = FC(\boldsymbol{p})$ is a contextual vertex of the full-correlation polytope* iff *the box $\boldsymbol{p}$ is a contextual vertex of the marginal n-cycle polytope. Moreover, $\boldsymbol{p}$ single marginals are given by $p_{o_i|I_i} = 1/2$.*

Proof. It follows from the definition that if $\boldsymbol{p}$ is a contextual vertex of the marginal polytope, $FC(\boldsymbol{p})$ is a contextual vertex of the full-correlation one. So we only need to prove that if $FC(\boldsymbol{p})$ is a vertex, then $\boldsymbol{p}$ is also a vertex.

Note that perfect (anti-)correlations demand some restrictions on the probabilities of $\boldsymbol{p}$, more precisely: perfect correlation, $\langle I_j I_{j+1} \rangle = 1$, implies that

$$p_{00|I_j I_{j+1}} + p_{11|I_j I_{j+1}} = 1 \text{ and } p_{01|I_j I_{j+1}} = p_{10|I_j I_{j+1}} = 0$$

[20] Please note that the condition of all correlators being positive implies that $\langle I_j \rangle = \langle I_k \rangle \forall j, k$

and perfect anticorrelation, $\langle I_j I_{j+1} \rangle = -1$, implies

$$p_{01|I_j I_{j+1}} + p_{10|I_j I_{j+1}} = 1, \text{ and, } p_{00|I_j I_{j+1}} = p_{11|I_j I_{j+1}} = 0.$$

For this scenario, the marginal conditions (equations (3.5)) are

$$p_{00|I_{j-1} I_j} + p_{10|I_{j-1} I_j} = p_{00|I_j I_{j+1}} + p_{01|I_j I_{j+1}},$$
$$p_{11|I_{j-1} I_j} + p_{01|I_{j-1} I_j} = p_{10|I_j I_{j+1}} + p_{11|I_j I_{j+1}}.$$

With these two conditions together we have

$$\langle I_{j-1} I_j \rangle = 1, \langle I_j I_{j+1} \rangle = 1 \implies p_{00|I_{j-1} I_j} = p_{00|I_j I_{j+1}} \text{ and } p_{11|I_{j-1} I_j} = p_{11|I_j I_{j+1}};$$
$$\langle I_{j-1} I_j \rangle = 1, \langle I_j I_{j+1} \rangle = -1 \implies p_{00|I_{j-1} I_j} = p_{01|I_j I_{j+1}} \text{ and } p_{11|I_{j-1} I_j} = p_{10|I_j I_{j+1}};$$
$$\langle I_{j-1} I_j \rangle = -1, \langle I_j I_{j+1} \rangle = 1 \implies p_{10|I_{j-1} I_j} = p_{00|I_j I_{j+1}} \text{ and } p_{01|I_{j-1} I_j} = p_{11|I_j I_{j+1}};$$
$$\langle I_{j-1} I_j \rangle = -1, \langle I_j I_{j+1} \rangle = -1 \implies p_{10|I_{j-1} I_j} = p_{01|I_j I_{j+1}} \text{ and } p_{01|I_{j-1} I_j} = p_{10|I_j I_{j+1}}.$$

Using the equations above recursively, we see that if a vertex of the full-correlation polytope has an odd numbers of correlators equal to -1, it is true that[21]

$$\langle I_j I_{j+1} \rangle = 1 \implies p_{00|I_j I_{j+1}} = p_{11|I_j I_{j+1}},$$
$$\langle I_j I_{j+1} \rangle = -1 \implies p_{01|I_j I_{j+1}} = p_{10|I_j I_{j+1}}.$$

So, each contextual vertex of the n-cycle full-correlation polytope uniquely determines one vertex of the non-full one. Also, we just need to invoke the normalisation condition for distributions to show that all single variable marginals are equal to $1/2$.

□

Lemma 5 (*n*-cycle inequalities are tight)**.** *The inequalities*

$$\sum_{j=1}^{n} \gamma_j \langle I_j I_{j+1} \rangle \leq n - 2, \tag{3.13}$$

with $\gamma_j \in \{-1, 1\}$, such that the number of negative coefficients γ_j is odd, represent facets of the noncontextual polytope.

Proof. By definition, an inequality represents a facet of the polytope of dimension d if it is saturated for (at least) d vertices of this polytope[22], and respected by all its elements.

Lemma 3 tell us that noncontextual vertices of the marginal polytope have an even number of correlators equal to -1, so they can never violate

[21] Since all contextual vertices are relabelling-equivalent to the one that $\langle I_1 I_2 \rangle = -1$, and all other correlators are equal to $+1$ (lemma 2), we could have restricted ourselves to prove this fact for this arrangement of correlators. In this case $\langle I_1 I_2 \rangle = -1$, $\langle I_2 I_3 \rangle = 1 \implies p_{01|I_1 I_2} = p_{11|I_2 I_2}$, the fact that all other correlators are positive implies in $p_{11|I_1 I_2} = p_{11|I_N I_1}$, and $\langle I_N I_1 \rangle = 1$, $\langle I_1 I_2 \rangle = -1 \implies p_{11|I_N I_1} = p_{10|I_1 I_2}$. So, $p_{01|I_1 I_2} = p_{10|I_1 I_2}$.

[22] Or equivalently, d elements that are affinely independent.

inequalities (3.13), and by convexity, this non-violation also holds for the entire noncontextual box.

Now, note that there are n ways to saturate an n-cycle inequality (3.13) with correlators $\langle I_j I_{j+1}\rangle = \pm 1$. These saturations happen when, for an specific k: $\langle I_k I_{k+1}\rangle = -\gamma_k$, for all other indexes[23] $j \neq k$: $\langle I_j I_{j+1}\rangle = \gamma_j$. Since there are two ways to obtain a list of even correlator with noncontextual vertices (lemma 3), there are $2n$ vertices of the marginal polytope that saturate inequalities (3.13).

It follows from the CH representation that the dimension of the noncontextual polytope is $2n$, so inequalities (3.13) represent facets. □

Lemma 6. *Let $\boldsymbol{C}$ be a contextual vertex of the marginal polytope of the n-cycle scenario. All its neighbouring vertices are noncontextual.*

Proof. First we will prove that if $\boldsymbol{C}_F^1$ and $\boldsymbol{C}_F^2$ are two different contextual vertices of the full-correlation polytope, they cannot be connected by edges of the full-correlation polytope. To do this, we will show that the uniform convex combination between these vertices,

$$\boldsymbol{p}_F := \frac{\boldsymbol{C}_F^1 + \boldsymbol{C}_F^2}{2},$$

is noncontextual by presenting two noncontextual vertices $\boldsymbol{NC}_F^1$ and $\boldsymbol{NC}_F^2$ such that

$$\boldsymbol{p}_F = \frac{\boldsymbol{NC}_F^1 + \boldsymbol{NC}_F^2}{2}$$

also holds.

Note that correlators of uniform convex combinations of two contextual vertices can assume three different values:

- $\langle I_j I_{j+1}\rangle_{\boldsymbol{p}_F} = 1$, when $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^1} = 1$ and $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^2} = 1$.
- $\langle I_j I_{j+1}\rangle_{\boldsymbol{p}_F} = -1$, when $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^1} = -1$ and $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^2} = -1$.
- $\langle I_j I_{j+1}\rangle_{\boldsymbol{p}_F} = 0$, when the correlators $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^1}$ and $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^2}$ have opposite signs.

Also, note that the number of correlators $\langle I_j I_{j+1}\rangle_{\boldsymbol{p}_F}$ that are equal to zero is even[24].

We will now show how to find the full-correlator boxes $\boldsymbol{NC}_F^1$ and $\boldsymbol{NC}_F^2$. For the sake of concreteness, we will first consider the case in which the number of correlators $\langle I_j I_{j+1}\rangle_{\boldsymbol{p}_F} = -1$ is even. Define the full-correlators $\langle I_j I_{j+1}\rangle_{\boldsymbol{NC}_F^1}$ and $\langle I_j I_{j+1}\rangle_{\boldsymbol{NC}_F^2}$ by the following rule:

[23] Please note that since there are an odd number of negative of coefficients γ_j, the choices of correlators that saturate (3.13) have an even number of negative coefficients γ_j.

[24] Define sets $\mathcal{C}^1$ and $\mathcal{C}^2$ by: $j \in \mathcal{C}^x$ if $\langle I_j I_{j+1}\rangle_{\boldsymbol{C}_F^x} = -1$. In this notation, the number of correlators $\langle I_j I_{j+1}\rangle_{\boldsymbol{p}_F}$ equal to zero is denoted by the number of elements in $\mathcal{C}^1 \Delta \mathcal{C}^2$, the symmetric difference between $\mathcal{C}^1$ and $\mathcal{C}^2$. Since $\#\mathcal{C}^x$ is odd, the relation $\#\mathcal{C}^1 + \#\mathcal{C}^1 = \#(\mathcal{C}^1 \Delta \mathcal{C}^2) + 2\#(\mathcal{C}^1 \cap \mathcal{C}^2)$ guarantees that $\#(\mathcal{C}^1 \Delta \mathcal{C}^2)$ is even.

- If $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = 1$, then $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^1} = \left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^2} = 1$;
- If $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = -1$, then $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^1} = \left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^2} = -1$;
- If $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = 0$, then $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^1} = 1$ and $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^2} = -1$.

Since the number of correlators $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F}$ equal to 0 or -1 is even, the two vertices $\boldsymbol{NC}_F^1$ and $\boldsymbol{NC}_F^2$ are in fact noncontextual. And follows that

$$\frac{\boldsymbol{C}_F^1 + \boldsymbol{C}_F^2}{2} = \frac{\boldsymbol{NC}_F^1 + \boldsymbol{NC}_F^2}{2}.$$

For the case in which the number of correlators $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = -1$ is odd, we define two full-correlator boxes, $\boldsymbol{NC}_F^1$ and $\boldsymbol{NC}_F^2$, as:

- If $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = 1$, then $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^1} = \left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^2} = 1$;
- If $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = -1$, then $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^1} = \left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^2} = -1$;
- If $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{p}_F} = 0$, then , for one specific index k we set $\langle I_k I_{k+1} \rangle_{\boldsymbol{NC}_F^1} = -1$ and $\langle I_k I_{k+1} \rangle_{\boldsymbol{NC}_F^2} = 1$. For all other indexes $j \neq k$ we set $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^1} = 1$ and $\left\langle I_j I_{j+1} \right\rangle_{\boldsymbol{NC}_F^2} = -1$.

It is easy to check that the number of negative correlators of the vertices $\boldsymbol{NC}_F^1$ and $\boldsymbol{NC}_F^2$ is even, and again the equivalence

$$\frac{\boldsymbol{C}_F^1 + \boldsymbol{C}_F^2}{2} = \frac{\boldsymbol{NC}_F^1 + \boldsymbol{NC}_F^2}{2}$$

holds.

To obtain the proof for the marginal polytope, we just need to add a simple trick to the previous proof technique. Again, we will show that if $\boldsymbol{C}^1$ and $\boldsymbol{C}^2$ are two contextual vertices of the marginal polytope, it is always possible to construct noncontextual boxes $\boldsymbol{NC}^1$ and $\boldsymbol{NC}^2$ such that

$$\frac{\boldsymbol{C}^1 + \boldsymbol{C}^2}{2} = \frac{\boldsymbol{NC}^1 + \boldsymbol{NC}^2}{2},$$

so, $\boldsymbol{C}^1$ and $\boldsymbol{C}^2$ cannot be neighbours.

From lemmas 1 and 4, we know that contextual vertices of the marginal polytope have an odd number of negative full-correlators and all single-correlators are equal to zero. So, we will show that it is always possible to construct noncontextual boxes that have the full-correlators of a noncontextual full-correlator vertex, but with single-correlators equal zero. In this way, we can explore exactly the same technique we used for the full-correlation case.

Let $\boldsymbol{NC^{1a}}$ and $\boldsymbol{NC^{1b}}$ be the two noncontextual vertices of the marginal polytope associated to the full-correlation vertex $\boldsymbol{NC_F^1}$ (lemma 3). The uniform convex combination

$$\boldsymbol{NC^1} := \frac{\boldsymbol{NC^{1a}} + \boldsymbol{NC^{1b}}}{2},$$

has the same full-correlators of $\boldsymbol{NC_F^1}$ and all single correlators equal zero[25].

[25]Just use the fact that the single correlators of $\boldsymbol{NC^{1a}}$ and $\boldsymbol{NC^{1b}}$ have opposite signs.

Analogously, we construct the noncontextual box

$$\boldsymbol{NC}^{\mathbf{2}} := \frac{\boldsymbol{NC}^{\mathbf{2a}} + \boldsymbol{NC}^{\mathbf{2b}}}{2},$$

that has the same full-correlators of $\boldsymbol{NC}^{\mathbf{2}}_{\boldsymbol{F}}$ and all single correlators equal to zero. So by the same arguments used on the full-correlation case, we can always find non contextual boxes $\boldsymbol{NC}^1$ and $\boldsymbol{NC}^2$ such that

$$\frac{\boldsymbol{C}^1 + \boldsymbol{C}^2}{2} = \frac{\boldsymbol{NC}^1 + \boldsymbol{NC}^2}{2}.$$

□

Before proving lemma 7, it is useful to define the concept of neighbour quasi-polytope of a vertex.

Definition 36 (Neighbour quasi-polytope). *Let $\mathcal{P}$ be a convex polytope with vertices $\mathcal{V} = \{V\}$. The neighbour quasi-polytope*[26] *associated to the vertex V is defined by*

$$N(V) := \{W \in \mathcal{P} \mid W = \frac{V + \alpha \sum_i \beta_i V_N^i}{1+\alpha}; \quad \alpha, \beta_i \geq 0, \sum_i \beta_1 = 1\}, \qquad (3.14)$$

where V_N^i are neighbour vertices[27] *of V.*

That is, the quasi-polytope of a vertex V is the set of all convex combinations that *necessarily* have non-null coefficient on V, and *may* have non-null coefficients on its neighbours V_N^i.

Lemma 7. *Let $\mathcal{P}$ be a convex polytope with vertices $\mathcal{V} = \{V\}$, and $\mathcal{S}(\mathcal{V})$ a set of vertices that are neighbour-free*[28] *on $\mathcal{P}$. The polytope $\mathcal{P}$ can be written as the disjoint union*

$$\mathcal{P} = \left(\bigcup_{V \in \mathcal{S}} N(V) \right) \cup conv(\mathcal{V} \setminus \mathcal{S}), \qquad (3.15)$$

where $N(V)$ is the neighbour quasi-polytope of V.

Proof. Since the set $\mathcal{S}(\mathcal{V})$ only contains vertices that are not connected by edges of $\mathcal{P}$, their neighbour quasi-polytopes are disjoint. Also, by definition, if $V \in \mathcal{S}$, then $\text{conv}(\mathcal{V} \setminus \mathcal{S}) \cap N(V) = \emptyset$.

To obtain the proof, we will use induction on the cardinality of the set $\mathcal{S}(\mathcal{V})$. For the case where the set $\mathcal{S}(\mathcal{V})$ has only a single vertex $\tilde{V}$, we need to prove that $\mathcal{P}$ can be written as the disjoint union

$$\mathcal{P} = \mathcal{P}^- \cup N(\tilde{V}), \qquad (3.16)$$

[26]Given a polytope $\mathcal{P}$ and a subpolytope $\mathcal{Q}$ of $\mathcal{P}$, we call $\mathcal{P} \setminus \mathcal{Q}$ a quasi-polytope. As a simple example, consider $\mathcal{Q}$ a facet or even a vertex of $\mathcal{P}$.

[27]Two vertices V and V_N are neighbour on $\mathcal{P}$ if they are connected by edges on $\mathcal{P}$.

[28]A set S of vertices is neighbour-free on a polytope $\mathcal{P}$ if no pair of vertices in $\mathcal{S}$ represents edges of $\mathcal{P}$.

with

$$\mathcal{P}^- := \text{conv}(\mathcal{V} \setminus \mathcal{S}).$$

For this, we have to show that all elements $R \in \mathcal{P}$ satisfy $R \in N(\tilde{V})$ and $R \notin \mathcal{P}^-$, or $R \in \mathcal{P}^-$ and $R \notin N(\tilde{V})$.

Every $R \in \mathcal{P}$ can be written as convex combination of the vertex $\tilde{V}$ and a point $Q \in \mathcal{P}^-$,

$$R = p\tilde{V} + (1-p)Q.$$

Let $\overline{\tilde{V}Q}$ be the segment between $\tilde{V}$ and Q,

$$\overline{\tilde{V}Q} := \{\lambda\tilde{V} + (1-\lambda)Q \mid \lambda \in [0,1]\},$$

that by convexity, intersects a facet of the polytope $\mathcal{P}^-$. In exactly one point[29] $Q' := \lambda'\tilde{V} + (1-\lambda')Q$. With the help of Q' we can separate all points $R \in \mathcal{P}$ in two classes:

- $p > \lambda'$, implies $R = p'\tilde{V} + (1-p')Q'$ for $p' \in (0,1]$, so R must lie in $N(V)$.

  To prove this we can check that $p'\tilde{V} + (1-p')Q' = R$ *iff* $p' = \frac{p-\lambda'}{1-\lambda'}$. So $p > \lambda'$ implies $p' > 0$.

- $p \leq \lambda'$, implies $R = p'Q' + (1-p')Q$ for $p' \in [0,1]$, so R must lie in $\mathcal{P}^-$.

  To prove this we can check that $p'Q' + (1-p')Q = R$ *iff* $p' = \frac{\lambda'-p}{1-\lambda'}$. So $p \leq \lambda'$ implies $p' \geq 0$.

So, we proved that the lemma holds when $\#\mathcal{S} = 1$. Now, we assume that the decomposition (3.15) holds for sets $\mathcal{S}$ of cardinality n and prove that this implies it holds for sets of cardinality $(n+1)$.

Define the set $\mathcal{S}^- := \mathcal{S} \setminus \{\tilde{V}\}$ that has all elements of $\mathcal{S}$, except for one specific vertex $\tilde{V} \in \mathcal{S}$, so $\#\mathcal{S} = \#\mathcal{S}^- + 1$. To finish the induction proof we have to show that if the disjoint union

$$\mathcal{P} = \left(\bigcup_{V \in \mathcal{S}^-} N(V)\right) \cup \text{conv}(\mathcal{V} \setminus \mathcal{S}^-), \tag{3.17}$$

is valid, then the disjoint union

$$\mathcal{P} = \left(\bigcup_{V \in \mathcal{S}^-} N(V)\right) \cup \left[\text{conv}(\mathcal{V} \setminus \mathcal{S}) \cup N(\tilde{V})\right], \tag{3.18}$$

is also valid.

Now, note that since the set $\mathcal{S}$ is neighbour-free on the polytope $\mathcal{P}$, we have the equivalence

$$\text{conv}(\mathcal{V} \setminus (\mathcal{S} \setminus \{\tilde{V}\})) = \text{conv}(\mathcal{V} \setminus \mathcal{S}) \cup N(\tilde{V}).$$

Now, substituting this equivalence in equation (3.17), we obtain the decomposition (3.18).

□

[29] Exactly on the facet of $\mathcal{P}^-$ that, if added to $N(V)$ would make it a polytope.

Lemma 8. *A contextual marginal box $\boldsymbol{p} \in \mathcal{M}(\mathcal{C}_n)$ violates one of the n-cycle inequalities (3.13).*

Proof. Exploring the fact that the contextual vertices of the n-cycle marginal polytope have an odd number of negative correlators and null single correlators, we can see that a box $\boldsymbol{p}$ respects the n-cycle inequalities *iff* it respects[30]

$$\boldsymbol{C} \cdot \boldsymbol{p} \leq n - 2,$$

for all contextual vertices $\boldsymbol{C}$.

Since the set of contextual vertices is neighbour-free on the marginal polytope (lemma 6), we can apply lemma 7 to write the marginal polytope as disjoint union of the noncontextual one and various contextual quasi-polytopes. This characterisation implies that any contextual box can be written as

$$\boldsymbol{p}_C = \frac{\boldsymbol{C} + \alpha \boldsymbol{C}_N}{1 + \alpha},$$

where the (noncontextual) box $\boldsymbol{C}_N$ satisfies[31] $\boldsymbol{C} \cdot \boldsymbol{C}_N = n - 2$. That is, all contextual points can be written in a convex combination that *necessarily* has one contextual vertex $\boldsymbol{C}$, and *may* have coefficients on a noncontextual box $\boldsymbol{C}_N$ that lies on the facet that "detects" $\boldsymbol{C}$. So,

$$\begin{aligned} \boldsymbol{C} \cdot \boldsymbol{p}_C &= \frac{\boldsymbol{C} \cdot \boldsymbol{C} + (1 - \alpha) \sum_i \beta_i \boldsymbol{C} \cdot \boldsymbol{N}_i^{\boldsymbol{C}}}{1 + \alpha} \\ &= \frac{n + \alpha(n-2)}{1 + \alpha} \\ &= n - \frac{2\alpha}{1 + \alpha}, \end{aligned}$$

which is larger than $(n - 2)$ for all $\alpha \geq 0$.

□

3.5.4 Quantum violations of the n-cycle inequalities

In this subsection we discuss what can be done with quantum boxes in the n-cycle scenario. We present the maximal violation that quantum boxes can attain of the n-cycle inequalities, and we see that $n \geq 4$ implies the existence of contextual quantum boxes.

Curiously, the problem of finding the maximum quantum violation of the n-cycle inequalities was solved *before* the complete characterisation of the noncontextual polytope. Although, focused on one particular scenario, Liang *et al* calculated the maximal violation for all inequalities with odd n [120] (see also [122]). And studying maximal quantum violations on Braunstein-Caves inequalities [91], Stephanie Wehner presented a proof that is general enough to apply to our case without modification [123].

Since the theorem is already proved in the above cited works, we will present it without a proof. We also remark that the proofs are constructive, in

[30] Here, $\boldsymbol{C} \cdot \boldsymbol{p}$ represents the canonical dot product.

[31] Please note that this point $\boldsymbol{C}_N$ corresponds to the point Q' used on the proof of lemma 7.

the sense that they present the quantum states and measurement operators that can be used to attain maximal violations.

Theorem 24. *Let $\boldsymbol{p}_Q$ be a quantum box on the n-cycle and $\boldsymbol{C}_k$ be a contextual vertex of the marginal polytope. The maximal quantum value of the n-cycle inequalities is given by*

$$\max \boldsymbol{p}_Q.\boldsymbol{C}_k = \begin{cases} \frac{3n\cos\left(\frac{\pi}{n}\right)-n}{1+\cos\left(\frac{\pi}{n}\right)} & \text{for odd } n, \\ n\cos\left(\frac{\pi}{n}\right) & \text{for even } n. \end{cases} \tag{3.19}$$

It is also possible to present some states and measurements that attain maximum quantum violation.

- For odd n, we can use the quantum state[32], $\rho := |\psi\rangle\langle\psi|$, and the measurement elements $I_j^0 = \left|I_j^0\right\rangle\left\langle I_j^0\right|$, $I_j^1 = I - \left|I_j^0\right\rangle\left\langle I_j^0\right|$, where $|\psi\rangle, \left|I_j^0\right\rangle$ are normalized vectors in $\mathbb{C}^3$,

$$|\psi\rangle := \begin{bmatrix}1\\0\\0\end{bmatrix}, \left|I_j^0\right\rangle := \begin{bmatrix}\cos\theta \\ \sin\theta\cos[j\pi(n-1)/n] \\ \sin\theta\sin[j\pi(n-1)/n]\end{bmatrix}, \tag{3.20}$$

  with $\cos^2\theta = \cos(\pi/n)/(1+\cos(\pi/n))$.

- For even n we can use the quantum state $\rho := |\psi\rangle\langle\psi|$, and the measurement elements: $I_j^0 = \left|I_j^0\right\rangle\left\langle I_j^0\right| \otimes \mathbb{1}$, $I_j^1 = \mathbb{1} - I_j^0$, even j ; $I_j^0 = \mathbb{1} \otimes \left|I_j^0\right\rangle\left\langle I_j^0\right|$, $I_j^1 = \mathbb{1} - I_j^0$ odd j, where

$$|\psi\rangle := \begin{bmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\\0\end{bmatrix}, \left|I_j^0\right\rangle := \frac{1}{N}\begin{bmatrix}1+\sin(j\pi/n)\\ \cos(\pi/n)\end{bmatrix}, \tag{3.21}$$

  with N being a normalization factor.

3.5.5 Future directions on the n-cycle

One of the most interesting aspects of the noncontextuality inequalities of the n-cycle is the fact that they are very simple. This could be useful to find the efficiency requirements necessary to attain a contextual violation in a non-perfect measurement scenario (see chapter 4).

A clear future direction is to understand what happens if we do not assume dichotomic inputs. For example, in the O-outcome 4-cycle, we know the CGLMP inequalities [90], that are tight, but do not characterise the whole noncontextual polytope [67].

Also, since the presented inequalities hold for all n, one could explore an asymptotic scenario, where n goes to infinity. We remark [124, 125], where the authors analyse the typical violation of quantum systems on the full-correlation N-partite scenario $(N, 2, 2)$.

[32] $|\psi\rangle\langle\psi|$ is the projector associated to $|\psi\rangle$.

Chapter 4

Physical implementation of nonlocal/contextual boxes

No creo en las brujas, pero que las hay las hay...

In the previous chapters, we have studied correlations on black boxes without worrying weather these boxes could be constructed in the real world. We avoided some natural questions: Are there nonlocal (resp. contextual) boxes in nature? Is it possible to construct nonlocal (contextual) boxes? Of course, if we believe in quantum mechanics, both answer are positive. But it would be interesting to construct a quantum nonlocal box, and guarantee its nonlocality only by analysing its input/output statistics, more on the direction of the device independence protocols discussed on section 2.9.

The existence of nonlocal boxes in real life has deep philosophical implications, for example, it forbids the interpretation that the results of any box outputs could be predicted with certainty, since we would need to know all variables that describe it. From a practical perspective, they provide perfect cryptography protocols and random number certification (see section 2.9).

We conclude this chapter by summarising the results presented in[1] [1, 2, 3], where we propose a physical implementation of a nonlocal quantum box.

Notation on quantum mechanics

In this chapter we will change our vocabulary a little bit, also, we will assume some familiarity with quantum mechanics and its standard notation. Readers who are not familiar with quantum mechanics can find a good introduction in [16, 126, 127, 128]. For now we just provide a brief introduction to the notation.

- Vectors lying in a vector space $\mathcal{H}$ will be denoted by $|\psi\rangle$, "ket psi", and its respective dual is defined by $\langle\psi| := \langle|\psi\rangle, .\rangle$, "bra psi". For

[1] Made in collaboration with Mateus Araújo, Adán Cabello, Daniel Cavalcanti, Marcelo Terra Cunha, Marcelo França, Jiří Minář, Valerio Scarani, and Colin Teo.

inner product between $|\psi\rangle, |\phi\rangle \in \mathcal{H}$, we use the shorthand "bra-ket" notation $\langle\psi|\phi\rangle := \langle|\psi\rangle, |\phi\rangle\rangle$. Please also note that we can represent linear operators with kets and bras, for example, if $|\psi\rangle$ is a normalised vector[2] $|\psi\rangle\langle\psi|$ is the projector onto the subspace spanned by $|\psi\rangle$.

- Quantum states that are represented by projectors, are called *pure states*. As usual in quantum mechanics we will refer to the pure state $\rho_\psi = |\psi\rangle\langle\psi|$ just as $|\psi\rangle$.

- The spectral theorem [129] states that all self-adjoint operators $A : \mathcal{H} \to \mathcal{H}$ admit the representation $A = \sum_a aA^a$ with $a \in \mathbb{R}$ and $A^a : \mathcal{H} \to \mathcal{H}$ being projectors that satisfies $A^a \geq 0$ and $\sum_a A^a = I$. Due to this fact, a self adjoint operator A is know in quantum mechanics as an *observable*, and the set $\{A^a\}$ is the measurement set associated to A.

- Pauli matrices, [130] $\sigma_i : \mathcal{H}_2 \to \mathcal{H}_2$, are unitary self-adjoint operators that are very useful in the study of quantum systems that lie in a bidimensional vector space. They are can defined as

$$\sigma_z := |g\rangle\langle g| - |e\rangle\langle e|, \quad \sigma_x := |g\rangle\langle e| + |e\rangle\langle g|, \quad \sigma_y := i|g\rangle\langle e| - i\,|e\rangle\langle g|, \tag{4.1}$$

  with $|g\rangle$ and $|e\rangle$ being an orthonormal basis.

4.1 Loophole free Bell tests

Physical experiments dedicated to construct nonlocal boxes are called *Bell tests*. Since 1972, quantum experimentalists have been reporting success on Bell tests [131, 132, 133, 134, 135, 136, 137], but as pointed out in [138], the "ultimate test of quantum mechanics" is still pending. This happens because all these experiments suffer from the so-called *loopholes*, strong assumptions that may not be well justified.

One main point is that in various statistical analysis, experimentalists implicitly adopt the *fair sampling* assumption. In real life experiments, it may happen that, in some rounds, the measurement results are not collected[3], so the experimentalists disconsider this uncollected data, and assume that what they have is a good representation of the whole. That is the fair sampling assumption, we will now illustrate how this assumption can fool the analysis of an experiment.

Alice observed a Joker flipping coins. She was able to count that, he tossed one specific coin 1000 times: 160 times she checked that heads was obtained, 40 times she checked that tails was obtained, and 800 times, for some reason, she missed the result.

After that, the Joker invited Alice for a head or tails bet. Due to her observation, she interfered that this is a biased coin: it outputs heads four times more than tails. Gladly, Alice accepted the bet and asked for heads.

[2]As usual in quantum mechanics, in this chapter, all our vectors $|\psi\rangle$ will be normalised.

[3]Or maybe the data is collected, but partially (or totally) corrupted, and for that reason, it is discarded.

What she did not know, is that 760 of the results that she missed turned out to output tails. And for this coin, the chances of obtaining tails after a toss is actually four times more likely than the chances of obtaining heads. Alice was fooled by believing that the sample that she had access was a "fair" representation of the whole.

From a fundamental point of view there is no reason to believe that nature maliciously disrespects fair sampling[4]. But this situation could be different in a cryptographic or device independent scenario, where one may be fighting against an active opponent who can exploit the fair sampling assumption to manipulate the results.

Another important point for obtaining a satisfactory Bell test is the *a priori non-signalling assumption*. This assumption is intricately related to the physical realisation of nonlocal/contextual boxes, and we need it to guarantee some interpretations on nonlocality[5] and the security of device independent[6] protocols. Before explaining it, we recall corollary 1 that states that all bipartite boxes admits the representation

$$p_{ab|A_xB_y} = \sum_{\lambda} \pi_{xy}(\lambda) p_{a|A_x;\lambda} p_{b|B_y;\lambda}$$

for some distributions $\pi_{xy} : \Lambda_{xy} \to [0,1]$, and one variable distributions $p_{.|A_x;\lambda}$ and $p_{.|B_y;\lambda}$. This theorem allows one undesired interpretation: it may happen that you have a nonlocal box, but its nonlocality arises from fact that a third part maliciously manipulates the result with the distributions π_{xy}. Now note that to construct such distributions, this third part needs to have access to Alice and Bob's inputs. The *a priori* non-signalling consists in forbidding that a third part can have access to Alice and Bob's inputs. For more on this interpretation, please see section 2.3, where we discuss how we can understand the probabilities of (local and nonlocal) multipartite boxes as a third party manipulation.

It happens that we cannot guarantee the *a priori* non-signalling only by analysing the statistical data of a box, but the security of device independent and cryptographic protocols relies on this assumption. One alternative to solve this problem is to assume a physical system that forbids a third part, to have access to all this information.

Special relativity is a well established physical theory which states that information takes a non-null time to propagate from one place to other. So, if we guarantee that all parties of a multipartite boxes are far apart and the time between their input choice and their output receiving is fast enough, we can guarantee the *a priori* non-signalling condition.

We remark that in some general box scenarios, we do not have the multipartite structure. Without it, we cannot invoke special relativity to ensure that a third part cannot have access to all input choices. For this reasons, the *a*

[4]Even though, scientific ethics states that one cannot claim properties of the whole only by analysing a fraction of the statistical data.

[5]Please see the discussions on theorem 16, that states that all multipartite boxes can be purified.

[6]That includes cryptographic protocols.

priori non-signalling assumption in contextuality tests may be very hard to justify.

The *a priori* non-signalling condition is guaranteed in [132, 134, 137], and fair-sampling independence in [135, 136], but these two conditions have never been simultaneously satisfied, it seems like they walk on opposite directions. We can ensure *a priori* non-signalling by exploring photonic systems[7], but they usually suffer from very inefficient measurements. Measurements on atomic systems can be efficient, but they cannot be performed fast enough to guarantee the *a priori* non-signalling condition.

It is widely believed that measurement inefficiencies on photonic system is merely a technological problem, so one way to obtain a loophole free Bell test is to wait for such technology and use polarisation measurements on photons [138]. Another approach to this problem is to seek for a physical setup that can be implemented in today's laboratories.

4.2 Measurement (in)efficiency models

In this section we will discuss how to deal with boxes that suffer from measurement inefficiencies. in the

4.2.1 The ideal box and the real box

In previous chapters, we dealt with boxes as a set of probability distributions. We will continue with this approach and with the same definition for boxes, but when looking for physical implementation of (nonlocal/contextual) boxes it is useful to introduce the abstraction of an "ideal box".

From a device independent perspective, the measured statistical data is the only relevant characteristic of a certain box. But for an experimentalist in a laboratory, the situation may be a little different. Any physical experiment is based upon some up to date theory that given a physical setup, can predict the probabilities of an abstract ideal box. Moreover, many times the experimentalists also know how to model the sources of noise and imperfections that prevent this ideal situation. With these informations, they are able to predict the statistical data that outputs from the "real box".

In this section we will explore methods that can be used to predict the statistical data of a real box by knowing its ideal behaviour and modelling imperfections. With it, we would like to guarantee that a certain physical setup can provide nonlocal/contextual correlations before actually implementing the real life experiment.

4.2.2 Missdetection

In a laboratory, Alice has a source that may emit a photon every ten seconds, and right in front the source, she puts a *photodetector*, an apparatus that "clicks" when a photon is absorbed. After some rounds, she noticed that sometimes

[7]Photons are useful because, assuming relativity, they are the fastest "information carriers" in nature.

her source emitted a photon, but for some reason the photodector did not click.

When a detector is supposed to click but it does not, we say that it *misdetects* a measurement[8]. Various measurement apparatus suffer from this kind of imperfections, and the misdetection happens with probability η, they can be modelled by

$$p_{\text{click}|\eta} = \eta p_{\text{click}}, \tag{4.2}$$

that induces

$$\begin{aligned} p_{\text{no-click}|\eta} &= 1 - p_{\text{click}|\eta} \\ &= 1 - \eta(1 - p_{\text{no-click}}), \end{aligned}$$

where p_{click} and $p_{\text{click}|\eta}$ can be understood as the probability of clicking in an ideal scenario and real scenario, respectively.

There is also a slightly different kind of misdetection. Suppose now Alice is measuring the linear polarisation of a photon. The possible outcomes of this measurements are vertical and horizontal, so the absence of a click is treated as a third outcome. These class of detectors can be modelled as

$$p_{H|\eta} := \eta p_H; \tag{4.3a}$$

$$p_{V|\eta} := \eta p_V; \tag{4.3b}$$

$$p_{\text{no-click}|\eta} := 1 - p_{H|\eta} - p_{H|\eta}, \tag{4.3c}$$

that admits a simple generalisation for various outcomes.

It important to remark that we can explore three outcome measurements in a two outcome scenario, we just map two different outcomes into one. For example, we can say that if the detector does not click, we will count as a click for the vertical polarisation[9].

We can also use equation (4.2) to analyse misdetection efficiencies in scenarios such that two or more measurement are jointly performed. For example, if the source emits two photons that are measured by two *independent* photodetectors with misdetection efficiency η_1 and η_2, we would write

$$p_{\text{click click}|\eta_1\eta_2} = \eta_1\eta_2 p_{\text{click click}}.$$

More abstractly, if A_x and B_y are jointly measured using independent detectors with misdetection efficiency η_x and η_y, we can connect the real probabilities with the ideal via[10]

$$p_{00|A_xB_y;\eta_x\eta_y} := \eta_x\eta_y p_{00|A_xB_y}; \tag{4.4a}$$

$$p_{01|A_xB_y;\eta_x\eta_y} := \eta_x p_{01|A_xB_y} + \eta_x(1-\eta_y)p_{00|A_xB_y}; \tag{4.4b}$$

$$p_{10|A_xB_y;\eta_x\eta_y} := \eta_y p_{10|A_xB_y} + \eta_y(1-\eta_x)p_{00|A_xB_y} \tag{4.4c}$$

$$p_{11|A_xB_y;\eta_x\eta_y} := 1 - p_{00|A_xB_y;\eta_x\eta_y} - p_{01|A_xB_y;\eta_x\eta_y} - p_{10|A_xB_y;\eta_x\eta_y}. \tag{4.4d}$$

[8]Another common source of error that can be understood as the opposite of misdetection is the *darkcount*. If a detector clicks when it is not supposed to with probability η_D, we say that it has a darkcount inefficiency η_D.

[9]For the CHSH scenario, Cyril [139] proved that we do not miss nonlocality by mapping the misdetection outcome to another one.

[10]Where we conveniently substitute the clicks for zeros and the absence of it by ones.

4.2.3 Quantum inefficiencies

In a certain laboratory, Alice has a source that produces a photon in the quantum state ρ and a detector that implements the measurement operator set $\{M^k\}$. In an ideal scenario, the probability to obtain a certain output k would be given by $p_{k|M;\rho} = \operatorname{tr} \rho M^k$, but in a real life laboratory, there can be various agents interacting with the photon between the source and the detector, and after such interactions, the probabilities of the results may differ from $\operatorname{tr} \rho M^k$.

Various physicists have dealt with this inefficiency problem, and have developed transformations[11] Φ that can be applied to the state ρ to obtain the probabilities in imperfect experiments via $p_{k|M;\rho,\Phi} = \operatorname{tr}(\Phi(\rho)M^k)$ [16, 140]. Here we will only focus on one of the most common imperfections on quantum systems, the *amplitude damping*.

We define the *amplitude damping* map with factor t on a quantum state $\rho : \mathcal{H} \to \mathcal{H}$ as

$$\Phi_t^{AD}(\rho) = \sum_k F(t)_k^* \rho F_k(t)$$

with $F_k^*(t)$ being the adjoint of

$$F_k(t) = \sum_{n=k} \sqrt{\binom{n}{k}} \sqrt{t^{n-k}(1-t)^k} |n-k\rangle\langle n|.$$

So if a quantum system suffers from an imperfection that can be modelled with amplitude damping, its probabilities will be given by $\operatorname{tr}(\Phi_t^{AD}(\rho)M_k)$.

Although we have introduced inefficiencies on quantum systems as a process that affects the quantum state, we could do it from a different but equivalent perspective. First, note that if we applied the adjoint of the amplitude damping map on the quantum measurement operators, we would have the same probabilities $\operatorname{tr}(\Phi_t^{AD}(\rho)M^k) = \operatorname{tr}(\rho \Phi_t^{*AD}(M_k))$. That is, we could say that there are no imperfections between the source and the detector, but the measurement apparatus has an amplitude damping imperfection. So, if inefficiencies of the detectors associated to the measurement set $\{M^k\}$ are modelled by amplitude damping, we can calculate its probabilities with $\{\Phi_t^{*AD}(M^k)\}$.

Although these two approaches may have different interpretations, both are equivalent, resembling the equivalence between Schrödinger and Heisenberg pictures [126].

4.3 Efficiency requirements for the CHSH scenario

In this section we analyse the efficiency requirements on a CHSH scenario that assumes the non-signalling conditions by proving a (slightly) more general version of theorems presented in [141, 142, 143, 139]. This result was obtained in collaboration with Mateus Araújo Santos.

[11] Also, in quantum mechanics we are only concerned with complete positive linear maps. For more on that we suggest [16].

Theorem 25. *Let $\boldsymbol{p} \in \mathcal{NS}(2,2,2)$ where measurements A_i and B_j have misdetection efficiencies η_{A_i} and η_{B_j}. The box $\boldsymbol{p}$ can be nonlocal iff one of the inequalities*

$$\begin{aligned}
&\eta_{A_0}\eta_{B_0} + \eta_{A_0}\eta_{B_1} + \eta_{A_1}\eta_{B_0} - \eta_{A_0} - \eta_{B_0} > 0; \\
&\eta_{A_0}\eta_{B_1} + \eta_{A_0}\eta_{B_0} + \eta_{A_1}\eta_{B_1} - \eta_{A_0} - \eta_{B_1} > 0; \\
&\eta_{A_1}\eta_{B_0} + \eta_{A_1}\eta_{B_1} + \eta_{A_0}\eta_{B_0} - \eta_{A_1} - \eta_{B_0} > 0; \\
&\eta_{A_1}\eta_{B_1} + \eta_{A_1}\eta_{B_0} + \eta_{A_0}\eta_{B_1} - \eta_{A_1} - \eta_{B_1} > 0;
\end{aligned} \tag{4.5}$$

holds[12].

Moreover, if one of the conditions (4.5) is satisfied, it is possible to find a quantum state and quantum measurements to construct a nonlocal quantum box in this misdetection scenario.

Proof. First, we recall that for perfect measurements, all non-trivial Bell inequalities are given by the CHSH inequalities, that are equivalent to the CH inequalities in the non-signalling regime (see subsection 2.4.2),

$$\begin{aligned}
-1 &\leq p_{00|A_0B_0} + p_{00|A_0B_1} + p_{00|A_1B_0} - p_{00|A_1B_1} - p_{0|A_0} - p_{0|B_0} \leq 0; \\
-1 &\leq p_{00|A_0B_1} + p_{00|A_0B_0} + p_{00|A_1B_1} - p_{00|A_1B_0} - p_{0|A_0} - p_{0|B_1} \leq 0; \\
-1 &\leq p_{00|A_1B_0} + p_{00|A_1B_1} + p_{00|A_0B_0} - p_{00|A_0B_1} - p_{0|A_1} - p_{0|B_0} \leq 0; \\
-1 &\leq p_{00|A_1B_1} + p_{00|A_1B_0} + p_{00|A_0B_1} - p_{00|A_0B_0} - p_{0|A_1} - p_{0|B_1} \leq 0.
\end{aligned}$$

The inequalities in the CH form are convenient due to the fact that they only have probabilities of obtaining the outcome 0, so we can use formula (4.4a) to write $p_{00|A_xB_y;\eta_x\eta_y} = \eta_x\eta_y p_{00|A_xB_y}$.

Define the CH parameter as,

$$CH := \eta_{A_0}\eta_{B_0}p_{00|A_0B_0} + \eta_{A_0}\eta_{B_1}p_{00|A_0B_1} + \eta_{A_1}\eta_{B_0}p_{00|A_1B_0} - \eta_{A_1}\eta_{B_1}p_{00|A_1B_1} - \eta_{A_0}p_{0|A_0} - \eta_{B_0}p_{0|B_0}.$$

Using the inequalities

$$-p_{00|A_1B_1} \leq 0, \quad p_{00|A_0B_1} \leq p_{0|A_0}, \quad \text{and} \quad p_{00|A_1B_0} \leq p_{0|B_0}$$

and factoring terms we can write

$$CH \leq \eta_{A_0}\eta_{B_0}p_{00|A_0B_0} - \eta_{A_0}(1-\eta_{B_1})p_{0|A_0} - \eta_{B_0}(1-\eta_{A_1})p_{0|B_0}.$$

Defining $p_{0|X_0} := \min\left(p_{0|A_0}, p_{0|B_0}\right)$ allows us to write

$$CH \leq \left(\eta_{A_0}\eta_{B_0} + \eta_{A_0}\eta_{B_1} + \eta_{A_1}\eta_{B_0} - \eta_{A_0} - \eta_{B_0}\right)p_{0|X_0}$$

so, if $\left(\eta_{A_0}\eta_{B_0} + \eta_{A_0}\eta_{B_1} + \eta_{A_1}\eta_{B_0} - \eta_{A_0} - \eta_{B_0}\right) \leq 0$, we have

$$CH \leq \left(\eta_{A_0}\eta_{B_0} + \eta_{A_0}\eta_{B_1} + \eta_{A_1}\eta_{B_0} - \eta_{A_0} - \eta_{B_0}\right)p_{0|X_0} \leq 0.$$

[12] Or equivalently, in this misdetection scenario, the set of non-signalling boxes $\mathcal{NS}(2,2,2)$ is strictly larger than the set of the local boxes $\mathcal{L}(2,2,2)$ iff one the inequalities 4.5 holds.

Due to the problem symmetry, we can obtain the other three inequalities presented in (4.5) via a simple relabelling argument. We finish the necessary condition proof by invoking a result obtained by Cyril [139] that guarantees that CH inequalities involving -1 require efficiency conditions that are stronger than the ones involving 0.

To prove that this condition is also sufficient, consider the state

$$|\psi\rangle = \frac{1}{\sqrt{1+\sin^2\alpha}}\Big(\cos\alpha|+\rangle_{A_1}|+\rangle_{B_1} + e^{i\varphi}\sin\alpha\Big(|+\rangle_{A_1}|-\rangle_{B_1} + |-\rangle_{A_1}|+\rangle_{B_1}\Big)\Big), \tag{4.6}$$

where $|\pm\rangle_{A_1}$ and $|\pm\rangle_{B_1}$ are the ±1 eigenstates of the A_1 and B_1 observables. We now define observables A_0 and B_0 by their eigenvectors as

$$|\pm\rangle_{A_0} := U_{A_0}|\pm\rangle_{A_1} \quad \text{and} \quad |\pm\rangle_{B_0} := U_{B_0}|\pm\rangle_{B_1},$$

where

$$U_{A_0} = U_{B_0} = \begin{pmatrix} \cos\alpha & e^{-i\varphi}\sin\alpha \\ e^{i\varphi}\sin\alpha & -\cos\alpha \end{pmatrix}.$$

The expected value of the CH operator constructed with these observables for $|\psi\rangle$ is

$$\langle CH\rangle_\psi = \frac{\sin^2\alpha}{1+\sin^2\alpha}\Big(\eta_{A_0}\eta_{B_0} + \eta_{A_0}\eta_{B_1} + \eta_{A_1}\eta_{B_0} - \eta_{A_0} - \eta_{B_0} - \eta_{A_0}\eta_{B_0}\sin^2\alpha\Big). \tag{4.7}$$

Since $\frac{\sin^2\alpha}{1+\sin^2\alpha} \geq 0$, if the inequality

$$\sin^2\alpha < \frac{1}{\eta_{A_0}\eta_{B_0}}\left(\eta_{A_0}\eta_{B_0} + \eta_{A_0}\eta_{B_1} + \eta_{A_1}\eta_{B_0} - \eta_{A_0} - \eta_{B_0}\right)$$

holds true for some α, we have $\langle CH\rangle_\psi > 0$. By hypothesis, the term in the parenthesis is positive, so we just need to guarantee that $\sin^2\alpha$ can be arbitrarily small, a fact that follows from the continuity of $\sin^2$. □

This general theorem allows us to recover some bounds as corollaries. For example, in [144], Ebehard found numerically that if all measurements have misdetection efficiencies η, it is possible to attain a quantum CHSH violation *iff* $\eta > 2/3$.

Corollary 4 (Ebehard bound[13]). *If $\eta_{A_0} = \eta_{B_0} = \eta_{A_1} = \eta_{B_1} = \eta$, we can attain a CHSH violation* iff

$$\eta > \frac{2}{3}. \tag{4.8}$$

Proof. In this case, inequalities (4.5) reduce to

$$\begin{aligned} 3\eta^2 - 2\eta &> 0; \\ \eta &> \frac{2}{3}. \end{aligned}$$

□

[13] This result was partially proved in [142] and completed in [139].

In [143], the authors analyse the case when Alice can perform measurements with efficiency η_A and Bob with efficiency η_B. We can now obtain their inequality by exploring theorem 25.

Corollary 5 (Larsson-Cabello bound[14])**.** *If $\eta_{A_0} = \eta_{A_1} = \eta_A$ and $\eta_{B_0} = \eta_{B_1} = \eta_B$, we can attain a CHSH violation* iff

$$\eta_A > \frac{\eta_B}{3\eta_B - 1}.$$

Proof. In this case, inequalities (4.5) reduce to

$$3\eta_A\eta_B - \eta_A - \eta_B > 0;$$
$$\eta_A > \frac{\eta_B}{3\eta_B - 1}.$$

□

Note that if $\eta_B = 1$, this bound reduces $\eta_A > 1/2$. This situation well approximates the case of Bell tests in which Bob performs measurements on atoms, that can be made very efficiently [135, 136].

In [141], G. Garbarino analyses the case in which each part can perform one measurement with missdeection efficiency η_0, and other with η_1. We can also recover his inequality with theorem 25.

Corollary 6 (Garbarino bound[15])**.** *If $\eta_{A_0} = \eta_{B_0} = \eta_0$ and $\eta_{A_1} = \eta_{B_1} = \eta_1$, we can attain a CHSH* iff

$$\eta_0 > 2(1 - \eta_1).$$

Proof. In this case, inequalities (4.5) reduce to

$$\eta_0^2 + 2\eta_0\eta_1 - 2\eta_0 > 0;$$
$$\eta_0 > 2(1 - \eta_1).$$

□

We call the attention to the fact that corollary 6 implies that if each part has access to one perfect detector ($\eta_1 = 1$), we just need non-null efficiency on the other detector to attain CHSH violation. This feature will be discussed again in section 4.4, where we present a photonic scenario in which we can attain CHSH violation with arbitrarily low photodetection efficiency.

4.4 A photonic proposal with quadrature measurements and photodetection

In this section we propose a photonic Bell test that, under some assumptions, can attain CHSH violation even with arbitrarily low photodetection efficiency. This proposal is obtained by exploring a scenario presented by Cavalcanti *et*

[14] This result was partially proved in [143] and completed in [139].
[15] This result was partially proved on [141].

al in [145] (see also [146, 147]) that combines photodetection and homodyne measurements on photonic quantum systems. The results presented in this section summarise the ones contained in [1, 2, 3].

In [1], we[16] develop an analytical approach to the scheme presented in [145], and with it, we show that it is possible to attain maximal quantum CHSH violation with the presented measurement operators. Also we analyse various states and their efficiency requirements to attain a loophole free Bell test in the scheme discussed.

In [2] we[17] prove that, assuming perfect transmittance, it is possible to attain CHSH violation even with arbitrarily low photodetection efficiency. Unfortunately, we also find some evidence that states with that property may not be viable for experimental implementation. We conclude that paper suggesting some possible directions for feasible implementations.

Still looking for loophole free Bell tests, in [3] we[18] propose an experimental setup to produce a class of quantum states that, in the photodetection/homodyne measurement scenario, can attain CHSH violation even with efficiency requirements that seem reasonable with current technology.

We hope that our findings can guide future research towards feasible proposals within the present Bell scenario.

4.4.1 Fock space

The photonic systems that we are going to explore on next subsections are described by an infinite dimensional vector space. We will write our quantum states and measurement operators using the *Fock space* representation. Here we will not worry about the physical motivation for that description , but adopt a practical approach: we define the set $\{|n\rangle\}_{i=0}^{\infty}$ as an orthonormal basis for $\mathcal{H}_\infty$ and allow the (informal) interpretation that $|n\rangle$ represents a quantum state of n photons.

4.4.2 Photodetection measurements

Photodetection is one of the most fundamental measurements on a photonic system. The intuition behind it is very simple, we have a detector that does not click for the vacuum state $|0\rangle$, and clicks with certainty for states orthogonal to it. For convenience[19], we assign the values -1 and $+1$ to a click and the absence of it, respectively.

In Fock basis, the ideal photodetection observable is represented by

$$D := |0\rangle\langle 0| - \sum_{n=1}^{\infty} |n\rangle\langle n|, \tag{4.9}$$

[16] Marco Túlio Quintino, Mateus Araújo, Daniel Cavalcanti, Marcelo França Santos, and Marcelo Terra Cunha.

[17] Mateus Araújo, Marco Túlio Quintino, Daniel Cavalcanti, Marcelo França Santos, Adán Cabello, and Marcelo Terra Cunha.

[18] Colin Teo, Mateus Araújo, Marco Túlio Quintino, Jiří Minář, Daniel Cavalcanti, Valerio Scarani, Marcelo Terra Cunha, and Marcelo França Santos.

[19] The motivation for this label is discussed on section 2.4.

and its measurement operators are

$$D^+ = |0\rangle\langle 0|, \quad D^- = \sum_{n=1}^{\infty} |n\rangle\langle n|.$$

4.4.3 Quadrature measurements

Another very important class of measurements in quantum optics is given by the *quadrature observables* [140]. For us, they are interesting because they may be efficiently implemented in a process called homodyne measurement [148]. If we recall that photonic systems are very good to close the locality loophole, homodyne measurements naturally arise as a good technique for loophole free Bell tests. Some schemes using homodyne measurements have been proposed, however, earlier results relying only on homodyne measurements required unfeasible setup and states [149, 150, 151, 152, 153] or displayed very small violations [154, 155], suggesting that homodyning may not render the definitive Bell test. In next subsections we will explore homodyne measurements from a different perspective, with it we would like to reinforce the idea that homodyne measurements may lead us to loophole free Bell tests.

The X quadrature observable can be defined as:

$$X := \int_{\mathbb{R}} x|x\rangle\langle x| \, \mathrm{d}x,$$

where $|x\rangle$ is defined in analogy with the quantum harmonic oscillator, and its inner product with an element of Fock basis is $\langle x|n\rangle = \varphi_n(x)$, with $\varphi_n(x)$ being the Hermite function[20] of order n [156]. We can represent the X quadrature observable in Fock basis,

$$X = \sum_{ij} \int_{\mathbb{R}} \varphi_i(x)\varphi_j(x) x |i\rangle\langle j| \, \mathrm{d}x.$$

One possible complication of quadrature measurements is that they can output an infinite number of outcomes, any $x \in \mathbb{R}$. But this issue can be solved with a *binning process*[21]. The most general binning is: we output the value $+1$ if the X measurement returns $x \in A^+$, where A^+ is a subset of the real numbers. We output the value -1 if $x \in A^- = \mathbb{R} \setminus A^+$.

With the projectors

$$Q^{\pm}(A^+) := \int_{A^{\pm}} |x\rangle\langle x| \, \mathrm{d}x,$$

we define the dichotomised version of the X observable as

$$Q(A^+) := Q^+(A^+) - Q^-(A^+)$$

[20] We remark the fact that the image of all Hermite functions is the set of real numbers $\mathbb{R}$.

[21] Binning [157] is a classical data processing technique to assign one certain value to various different outputs. For example, one natural binning consists in assigning the value $+1$ to all outcomes that are larger than zero and -1 to all outcomes smaller than zero. We remark that, given its classical description, one cannot use the binning process to transform local statistical data into nonlocal one.

so our dichotomised quadrature measurement operators are

$$Q^+(A^+) = \sum_{ij} \int_{A^+} \varphi_i \varphi_j |i\rangle\langle j|, \quad Q^-(A^+) = \sum_{ij} \int_{A^-} \varphi_i \varphi_j |i\rangle\langle j|.$$

We now call the attention to the fact that the X operator is unbounded [158]. And as we will see in subsection 4.4.9, unbounded operators may bring some unexpected physical issues.

4.4.4 A convenient subspace

Since[22] $Q^2 = I$, the subspace generated by $\{|0\rangle, Q|0\rangle\}$ is invariant under Q. This allows us to write

$$Q|0\rangle = \cos\theta|0\rangle + \sin\theta|\Xi\rangle, \tag{4.10}$$

with

$$\big|\Xi(A^+)\big\rangle := \frac{2}{\sin\theta} \sum_{n=1}^{\infty} \int_{A^+} \varphi_0(x)\varphi_n(x)\,\mathrm{d}x|n\rangle \tag{4.11}$$

being a normalised vector orthogonal to $|0\rangle$ and[23]

$$\cos\theta(A^+) := 2\int_{A^+} \varphi_0(x)^2\,\mathrm{d}x - 1. \tag{4.12}$$

Moreover, in this subspace, the operators D and Q have a very useful representation.

Lemma 9. *D and Q can be written as*

$$D = \Pi D \Pi + (\mathbb{1} - \Pi)D(\mathbb{1} - \Pi)$$
$$Q = \Pi Q \Pi + (\mathbb{1} - \Pi)Q(\mathbb{1} - \Pi),$$

where Π is the projector onto the subspace spanned by $\{|0\rangle, Q|0\rangle\}$.

Proof. Note that $\{|0\rangle, Q|0\rangle\}$ is an invariant subspace of both operators D and Q, as

$$Q(\alpha|0\rangle + \beta Q|0\rangle) = \alpha Q|0\rangle + \beta|0\rangle$$

and

$$\begin{aligned} D(\alpha|0\rangle + \beta Q|0\rangle) &= \alpha D|0\rangle + \beta DQ|0\rangle \\ &= \alpha|0\rangle + \beta(2|0\rangle\langle 0| - I)Q|0\rangle \\ &= \alpha|0\rangle + \beta(2|0\rangle\langle 0|Q|0\rangle - Q|0\rangle) \\ &= (\alpha - 2\beta\langle 0|Q|0\rangle)|0\rangle + \beta Q|0\rangle. \end{aligned}$$

Since both D and Q are self-adjoint, it follows that the pre-image of $\{|0\rangle, Q|0\rangle\}$ is also within $\{|0\rangle, Q|0\rangle\}$, so the orthogonal decomposition is valid for both operators. □

[22] The dichotomised X quadrature observable is always a function of the binning set A^+, but if the binning has minor importance, we will just write Q.

[23] Please note that one binning set A^+ defines θ, but the same θ can be attained with various different binnings.

It is important to note that $|\Xi\rangle$ is an eigenvector of D. With this notation, the restriction of Q and D to this subspace, in the orthonormal basis $\{|0\rangle, |\Xi\rangle\}$, takes the form

$$\begin{aligned} Q_R &:= \Pi Q \Pi = \cos\theta\ \sigma_z + \sin\theta\ \sigma_x, \\ D_R &:= \Pi D \Pi = \sigma_z. \end{aligned}$$

4.4.5 CHSH scenario with photodetection and X quadrature measurements

In [145] , Cavalcanti *et al.* explore a Bell test scheme that combines photodetection and homodyne measurements and show that it is possible to attain a CHSH violation of 2.25 with the state $|\psi\rangle = \frac{|02\rangle+|20\rangle}{\sqrt{2}}$. Also their scheme can tolerate some inefficiencies: photodetection $\eta = 0.71$, or transmission $t = 0.84$ (for discussion of photonic inefficiencies, please read subsection 4.2.2), efficiency requirements that are comparable to the ones obtained for polarisation Bell test schemes [142]. After presenting some other examples they left open the maximal amount of violation for other states.

For this scenario, it is useful to define the CHSH operator

$$CHSH_P(A^+) := D \otimes D + D \otimes Q(A^+) + D \otimes D - Q(A^+) \otimes Q(A^+),$$

and with an analytical approach, we explicitly calculate the maximum CHSH violation as a function of the binning choice A^+, and also provide quantum states that attain it.

Theorem 26. *The maximal violation of $CHSH_P(A^+)$ as a function of the binning choice is given by*

$$\|CHSH\|_P(A^+) = 2\sqrt{1+\sin^2\theta(A^+)}.$$

Moreover, it is always possible to find a quantum state $|\psi\rangle$ that attains such violation by diagonalizing the 4×4 matrix that represents $CHSH_P$ restricted to the subspace spanned by $\{|0\rangle, Q|0\rangle\}^{\otimes 2}$

Proof. Invoking the Tsirelson-Khalfin-Landau identity (equation (2.7)), we see that

$$\|CHSH\|_P^2 = 4 + \|[Q,D]\|^2.$$

Using lemma 9 one can check that $[Q,D] = [\Pi Q \Pi, \Pi D \Pi]$, where $\Pi : \mathcal{H}_\infty \to \mathcal{H}_2$ is the orthogonal projector onto the subspace spanned by $\{|0\rangle, Q|0\rangle\}$. With it we can explicitly calculate the commutator

$$\begin{aligned} [Q,D] &= \cos\theta\ [\sigma_z,\sigma_z] + \sin\theta\ [\sigma_x,\sigma_z] \\ &= 0 - \sin\theta\ 2i\sigma_y, \end{aligned}$$

and use the fact that $\|\sigma_y\| = 1$ to finish the proof. □

In the next subsections, we analyse this scenario under measurement imperfections.

4.4.6 An imperfect measurement scenario

A photodetector has efficiency η if the probability of outputting $+1$ for the state $|n\rangle$ is[24] $1-(1-\eta)^n$, that is, the complement of the probability that the apparatus misdetects all photons. This is equivalent to saying that we applied an amplitude damping map with factor η on the photodetection measurement operators [140] (see section 4.2.3). For analysing imperfect photodetectors we define

$$D_\eta^- := \sum_{n=1}^{\infty}(1-(1-\eta)^n)|n\rangle\langle n|, \tag{4.13a}$$

$$D_\eta^+ := |0\rangle\langle 0| + \sum_{n=1}^{\infty}(1-\eta)^n|n\rangle\langle n|, \tag{4.13b}$$

that induces the observable $D_\eta := D_\eta^+ - D_\eta^-$. Please note that if we restrict ourselves to states lying in the subspace spanned by $\{|0\rangle, |\Xi\rangle\}$, photodetection measurements can be described by

$$D_{R\eta}^+ := \left\langle \Xi \middle| D_\eta^+ \middle| \Xi \right\rangle |\Xi\rangle\langle\Xi|, \qquad D_{R\eta}^- := I - D_{R\eta}^+$$

so, for all quantum states that lie in the subspace spanned by $\{|0\rangle, |\Xi\rangle\}$ we have $p_{+1|D_\eta} = \left\langle \Xi \middle| D_\eta^+ \middle| \Xi \right\rangle p_{+1|D}$. This identity motivates us to define the overall photodetection efficiency

$$H : [0,1] \to [0,1], \qquad H(\eta) := \left\langle \Xi \middle| D_\eta^+ \middle| \Xi \right\rangle,$$

and with it, we can make a clear analogy with the efficiency model presented on equation (4.4a).

Another source of errors is the transmittance t which affects both measurements: photodetection and homodyning and it is also modelled by the amplitude damping map [140, 16]. For the photodetection, the effects of transmittance are equivalent to photodetection efficiency[25], so we define

$$D_{R\eta t}^+ := H(\eta t)|\Xi\rangle\langle\Xi|, \quad D_{\eta t}^- := I - D_{R\eta t}^+,$$

that can be used to calculate probabilities for states that lie in $\mathrm{span}(\{|0\rangle, |\Xi\rangle\})$.

Unfortunately, quadrature measurements with transmittance t do not admit such a simple representation, so we represent them by[26]

$$Q_t^+ = \sum_k F_k^* Q^+ F_k, \quad Q_t^- = \sum_k F_k^* Q^- F_k,$$

with

$$F_k = \sum_{n=k} \sqrt{\binom{n}{k}} \sqrt{t^{n-k}(1-t)^k} |n-k\rangle\langle n|.$$

---

[24] Adopting the convention that $0^0 = 1$.

[25] We also assume that the sources of errors are uncorrelated.

[26] See subsection 4.2.3.

4.4.7 Violation for all non-null η photodetection efficiency

Now we can analyse a more realistic quantum Bell test by using the operator

$$CHSH_{P\eta t}(A^+) := D_{\eta t} \otimes D_{\eta t} + D_{\eta t} \otimes Q_t(A^+) + D_{\eta t} \otimes D_{\eta t} - Q_t(A^+) \otimes Q_t(A^+).$$

Note that if $t = 1$, corollary 6 states that we can verity quantum nonlocality with *arbitrarily low* photodetection efficiency. Furthermore, by projecting Q_t, $D_{\eta t}$ as before, we can attain the required violation with states lying in span$(\{|0\rangle, |\Xi\rangle\}^{\otimes}2)$. Moreover, the demonstration of theorem 25 informs us that the state

$$|P_H\rangle := \frac{1}{\sqrt{1+\sin^2\frac{\theta}{2}}}\left[\cos\frac{\theta}{2}|{+}{+}\rangle + \sin\frac{\theta}{2}(|{+}{-}\rangle + |{-}{+}\rangle)\right],$$

with $|\pm\rangle$ being the normalised eigenvectors of Q_R, attains

$$\left\langle CHSH_{\eta 1P}\right\rangle_{P_H} = 2 + H^2(\eta)\frac{4\cos^2\frac{\theta}{2}\sin^4\frac{\theta}{2}}{1+\sin^2\frac{\theta}{2}},$$

which is larger than 2 for every $\eta > 0$ and nontrivial choice of θ.

Also, the critical transmittance[27] of these states is ≈ 0.92, and is reached in the limit $\theta \to 0$. We have also maximized $\left\langle CHSH_{\eta 1P}\right\rangle$ and found that the binning choice that reaches the largest violation obeys $\cos\frac{\theta}{2} = (\sqrt{5}-1)/2$.

Also, since in the subspace span$(\{|0\rangle, |\Xi\rangle\})$, the operator $CHSH_{\eta tPR}$ is just a 4×4 matrix, and we can find the eigenvector $|P_\Xi\rangle$ that provides us the smaller requirements on efficiency[28]. The minimum requirements on efficiency for the states $|P_H\rangle$ and $|P_\Xi\rangle$ are presented on figure 4.1.

Before ending this subsection, we remark that in [1] we analyse the behaviour of $\left\langle CHSH_{P\eta t}(A^+)\right\rangle$ for various different classes of states, we also present some specific examples and considered *dark counts* inefficiencies[29].

4.4.8 An atom-photon scenario

Now we are going to analyse scenarios in which Alice performs measurements on atomic observables, described by Pauli matrices, and Bob measures Q or D. This scheme was explored in ref. [159], and it is interesting because measurement efficiencies on atomic systems can be very high [135, 136], so we can tolerate larger inefficiencies on photonic measurements (see lemma 5).

Define the CHSH operator

$$CHSH_{A\eta t} := V(\gamma) \otimes D_{\eta t} + V(\gamma) \otimes Q_t + V(-\gamma) \otimes D_{\eta t} - V(-\gamma) \otimes Q_t,$$

---

[27] Here, critical transmittance is the minimal transmittance required for attaining a CHSH violation.

[28] Please note that $|P_\Xi\rangle$ is an eigenvector of the restricted operator $CHSH_{\eta tPR}$, not of $CHSH_{\eta tP}$. In a non-perfect measurement scenario, we do not guarantee that the quantum states that attain maximal violation lie inside span$\{|0\rangle, |\Xi\rangle\}^{\otimes 2}$, that is, these are not the minimum efficiency requirements for this setup.

[29] Dark counts can be seen as the opposite of the photodetection efficiency, instead of misdetecting a photon with probability η, the apparatus misdetect the vacuum with probability η_D.

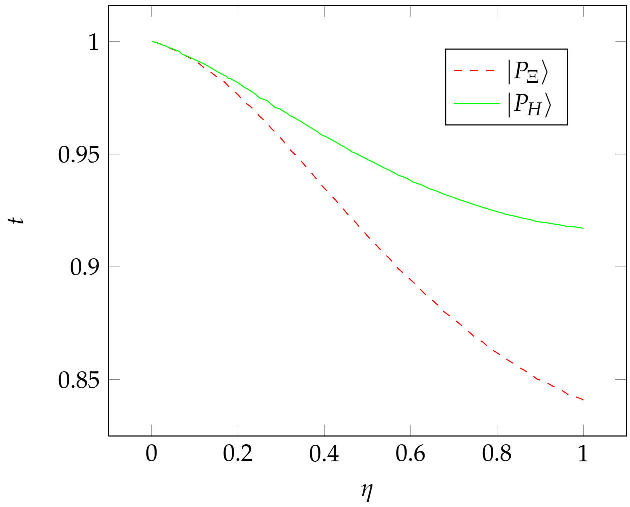

Figure 4.1: Contour lines for $\langle CHSH_{P\eta t}\rangle = 2$, considering the efficiency η of the photodetectors and the transmittance t between the source and the photodetectors.

where[30]

$$V(\gamma) := \cos\gamma\,\sigma_z + \sin\gamma\,\sigma_x.$$

In order to obtain an analytical result, we now restrict ourselves to the subspace spanned by $\{|g\rangle, |s\rangle\} \otimes \{|0\rangle, |\Xi\rangle\}$, where $|g\rangle$ and $|s\rangle$ are two atomic levels, eigenstates of the σ_z observable. Also, we assume ideal transmittance ($t = 1$) and, for simplicity, we only consider binnings A^+ such that $\int_{A^+} \varphi_0(x)^2\,\mathrm{d}x = 1/2$, where $Q_R = \sigma_x$.

In this case, the Bell operator $CHSH_{A\eta t}$ is represented by a 4×4 matrix with eigenvector

$$|A_\Xi\rangle \propto \left[(1-H)\cos\gamma + \sqrt{\sin^2\gamma + (1-H)^2\cos^2\gamma}\right]|g\,0\rangle + \sin\gamma|s\,\Xi\rangle,$$

that attains

$$\langle CHSH_A\rangle_{A_\Xi} = 2H\cos\gamma + 2\sqrt{\sin^2\gamma + (1-H)^2\cos^2\gamma}, \qquad (4.14)$$

which is larger than 2 if $\eta > 0$ and $\gamma \in \left(0, \frac{\pi}{2}\right)$. That is, we have a violation of a Bell inequality with *arbitrarily low photodetection efficiency* for all choices of non-commuting atomic observables.

We also compute numerically the efficiency requirements for $\langle CHSH_{A\eta t}\rangle_{A_\Xi} > 2$, result presented on figure 4.2.

[30] Since we have a two level quantum system and will use two self-adjoint operators with eigenvalues ± 1, they can be written $M = |0_m\rangle\langle 0_m| - |1_m\rangle\langle 1_m|$ and $N = |0_n\rangle\langle 0_n| - |1_n\rangle\langle 1_n|$. For noncommuting M and N, there is only one $\gamma \in \left(0, \frac{\pi}{2}\right)$ defining $2\cos\frac{\gamma}{2}|a\rangle := |0_m\rangle + |0_n\rangle$, $2\sin\frac{\gamma}{2}|b\rangle := |0_m\rangle - |0_n\rangle$. In this convenient basis, we can write $V(\gamma) = |a\rangle\langle a| + |b\rangle\langle b|$ so that $M = V(\gamma)$ and $N = V(-\gamma)$, and one can interpret 2γ as the angle between the two observables.

4.4.9 Looking for feasible states

Although the state $|A_\Xi\rangle$ could be useful for loophole free Bell tests, we do not know any experimental setup to produce it. Moreover, we have strong evidence that the states $|\Xi(A^+)\rangle$ may not be practical for physical implementations. For the binning $A^+ = [0,\infty]$, we can analytically solve the integrals to write

$$|\Xi([0,\infty])\rangle = 2\sum_{n=0}^{\infty} \frac{(-1)^n(2n)!}{\sqrt{2\pi(2n+1)!\,2^n n!}}|2n+1\rangle,$$

that asymptotically, assumes the form

$$|\Xi([0,\infty])\rangle \sim \frac{2^{3/4}}{\sqrt{\zeta\left(\frac{3}{2}\right)\left(2\sqrt{2}-1\right)}} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^{\frac{3}{4}}}|2n+1\rangle,$$

that can be used to check that we have a divergence on the expected value of the photon-number observable $N := \sum_n n|n\rangle\langle n|$,

$$\langle N\rangle_{|\Xi([0,\infty])\rangle} = \langle\Xi|N|\Xi\rangle =\propto \sum_n \frac{1}{\sqrt{n}} \to \infty.$$

Physically, this divergence also implies that to produce this state one would need an infinite amount of energy. We also have numerical evidence that this photon-number divergence occurs for *all* binning choices A^+, indicating that states lying in the subspace $\text{span}\{|0\rangle, |\Xi\rangle\}^{\otimes 2}$ are not feasible.

In order to find a feasible Bell test, it is physically sound to replace $|\Xi\rangle$ with a more familiar state. A good candidate is the so-called odd cat state [160], whose production has been explored experimentally [161],

$$|\text{cat}\rangle := \frac{|\alpha\rangle + |-\alpha\rangle}{\sqrt{2}\sqrt{1+e^{-2|\alpha|^2}}},$$

with $|\alpha\rangle$ being the coherent state [162].

$$|\alpha\rangle := e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle.$$

With the cat state, we can define [163]

$$|A_{\text{cat}}\rangle = \cos\nu|g\,0\rangle + \sin\nu|e\,\text{cat}\rangle,$$

that does not attain violation for all non-null photodetection efficiencies but still provides good numbers: for example, in a perfect measurement scenario, it attains[31] $\langle CHSH_{A\eta t}\rangle_{A_{\text{cat}}} = 2.60$. The efficiency requirement for $|A_{\text{cat}}\rangle$ is shown in figure 4.2.

Another idea is to replace $|\Xi\rangle$ for the coherent state $|\alpha\rangle$ to define

$$|A_\alpha\rangle = \cos\nu|g0\rangle + \sin\nu|e\alpha\rangle,$$

[31] $\alpha \approx 2.20i$, $\gamma \approx 0.59$, $\nu \approx 0.74$, and $A^+ \approx [-0.55, 0.55]$.

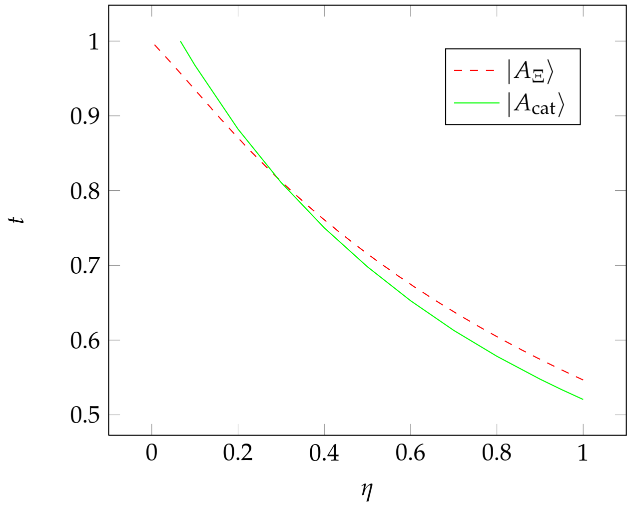

Figure 4.2: Contour line on which $\langle CHSH_A \rangle = 2$, considering the efficiency η of the photodetectors and the transmittance t between the source and the photodetectors. For $|A_{\text{cat}}\rangle$ violations can be obtained above critical efficiency 0.066, for which $\alpha \approx 2.29i$, and above critical transmittance ≈ 0.52, reached for $\alpha \approx 2.87i$. For $|A_\Xi\rangle$ the critical efficiency is zero, while the critical transmittance is ≈ 0.55.

that, in a perfect measurement setup, attains[32] $\langle CHSH_{A11}\rangle_{\psi_\alpha} = 2.32$. Moreover, it is possible to attain a CHSH violation even for[33] $\eta = 0.15$, or transmittance[34] $t = 0.55$. Also, for a detection efficiency of $\eta = 0.8$ and a transmission of $t = 0.8$ we find[35] $\langle CHSH_{A\eta t}\rangle_{A_\alpha} = 2.07$. In figure 4.3, we present the efficiency requirements for $|A_\alpha\rangle$. For comparison, we have also included the curve $\eta t = 2/3$ that results from the Eberhard bound [144, 142], which is the required efficiency and transmission to perform a loophole-free experiment with photon polarisation, and the curve from the best experimental proposal to date involving an atom and a photonic mode [159].

4.4.10 An experimental proposal

Although some papers reported feasibility of the states $|A_{\text{cat}}\rangle$ and $|A_\alpha\rangle$ MT: aguardando nome dos trabalhos, pedi para o frança: micromaser do Haroche , we do not know how to adapt the their physical state preparation to our measurement scheme. Seeking for a viable Bell test, in [3] we propose a physical setup to construct a state ρ_α, which is equivalent to $|A_\alpha\rangle$ in some regimes. Assuming the numbers found in recent papers [164, 165], ρ_α does not attain the same good results as $|A_\alpha\rangle$, but still provides reasonable CHSH

[32] $\alpha \approx 2.10i$, $\gamma \approx 0.55$, $\nu \approx 0.77$, and $A^+ \approx [-0.53, 0.53]$.
[33] $\alpha \approx 3.35i$, $\gamma \approx 0.14$, $\nu \approx 0.16$, and $A^+ \approx [-0.34, 0.34]$.
[34] $\alpha \approx 3.38i$, $\gamma \approx 0.03$, $\nu \approx 0.33$, and $A^+ \approx [-0.44, 0.44]$.
[35] $\alpha \approx 2.33i$, $\gamma \approx 0.34$, $\nu \approx 0.66$, and $A^+ \approx [-0.53, 0.53]$.

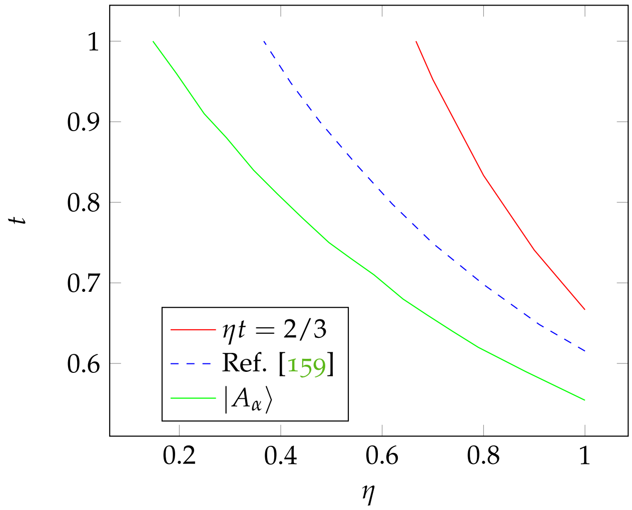

Figure 4.3: Contour line for $\langle CHSH\rangle_{A\eta t} = 2$. The green line (bottom) corresponds to the ideal state of our proposal. The parameters α, γ, ν were optimized for each point. For comparison, we include the curve $\eta t = 2/3$ (red, on top) that results from the Eberhard bound [144, 142], and the curve from the best experimental proposal to date [159] (blue, dashed).

violation with low requirements on transmittance t and and photodetection efficiency η.

Since methods to construct quantum states and current technology on experimental optics are not on the scope of this dissertation, we just present the form of state ρ_α (equation (4.15)) and some contour lines (figure 4.4) for $\left\langle CHSH_{A\eta t}\right\rangle_{\rho_\alpha}$ for the cases our state can be produced by adapting the scheme presented in [164, 165]. For more details, please see [3].

The scheme proposed construct the state[36]

$$\rho_\alpha = V|A_\alpha\rangle\langle A_\alpha| + (1-V)\sigma, \tag{4.15}$$

where

$$\begin{aligned} |A_\alpha\rangle &= \cos\nu|s0\rangle + \sin\nu|g\alpha\rangle; \\ \sigma &= \cos^2\nu|s0\rangle\langle s0| + \sin^2\nu|g\alpha\rangle\langle g\alpha|; \\ V &= \exp\Big[-\left(\frac{\kappa_b}{\kappa_c} + |r_{BS}|^2\right)\frac{|\alpha|^2}{2|t_{BS}|^2}\Big]. \end{aligned}$$

The parameters r_{BS}, t_{BS} stand for reflectance and transmittance of a beam splitter used in the process of generating the state, κ_b and κ_c stand for the decay rate of the mirrors of a cavity with an atom in the dispersive regime. We remark that it is possible to construct beam splitters with any reflectance and transmittance that respects $|r_{BS}|^2 + |t_{BS}|^2 \leq 1$, and cavities with decay rates satisfying $\kappa_b/\kappa_c = 1/20$ have been reported on [164].

[36] That can be understood as a convex combination of the ideal state with its "depolarised" version.

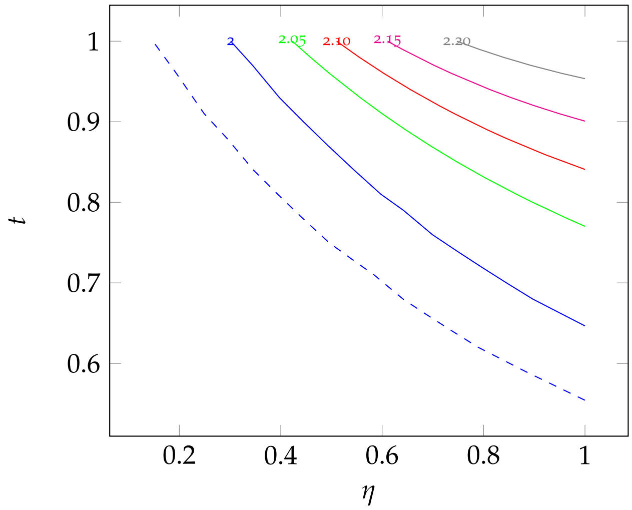

Figure 4.4: Contour lines of $\langle CHSH_{A\eta t}\rangle_{\rho_\alpha}$ as a function of η and t. We keep $|\alpha| = 2$ fixed and optimize the measurements and state parameters γ, ν for each point. The experimental parameters used are $\kappa_b/\kappa_c = 1/20$ and $|r_{BS}|^2 = 0.01$. The critical line for $|A_\alpha\rangle$ (dotted line) is included for comparison.

4.4.11 Possible future directions

One clear future direction is to look for laboratories that can implement the physical system proposed. And also to look for restrictions and limitations of some specific laboratories to optimise our scenario for their setups.

It is also natural to look for scenarios in which the photonic part explores homodyning measurements on the dichotomised version of different quadratures [140]. If we could attain reasonable CHSH violation only with homodyne measurements we would not need to worry about photodetection efficiencies anymore.

Another direction would be to improve our results by adapting our scheme to Bell scenarios with more parties and inputs. We remark that references [166, 167] show examples of Bell scenarios that are less demanding on efficiencies than the CHSH, and [168] explores dichotomised homodyne measurements in a multipartite scenario.

Conclusions and perspectives

In this dissertation we developed a black box framework that can be used to infer properties of physical experiments just by analysing its statistical data. We introduced the concepts of signalisation, locality, and noncontextuality, discussing their interpretations, proving theorems, and presenting examples. In particular a box is nonlocal/contextual if its probabilities cannot be understood as ignorance on hidden variables which correlate its subsystems.

After presenting various known results on multipartite boxes, we developed a more general concept of black boxes, that allows us to study any set of probability distributions. In this general box formalism, we explored some results on convex geometry to analyse the n-cycle scenario, obtaining (for the first time) a simple halfspace characterisation for a class of infinitely many noncontextual polytopes.

We discussed the meaning of the physical realisation of nonlocal and contextual boxes, and why performing these experiments is so hard in real laboratories. In order to understand the difficulties of constructing nonlocal boxes in a imperfect scenario, we proved some efficiency requirements to attain a loophole free Bell test. These efficiency bounds motivated us to look for physical setups that combine very efficient measurements with not so efficient ones. Pairing homodyne measurements (that can be made very efficient) with photodection (usually, not very efficient) we presented an optical setup that can attain CHSH violation even for detectors with arbitrarily low photodetection efficiency. But we concluded that the quantum states that attain such numbers may not be physically viable.

Still in the same homodyne/photodetection scheme, we presented states that can be constructed by adapting some state of the art physical setups. These states do not attain the same impressive results as the (possibly) non-feasible ones, but still provide nonlocal statistics even with efficiency requirements that seem reasonable with current technology.

Although various questions were answered, we finish this text with some open ones. What more can we learn from that simple characterisation of the noncontextual polytope for the n-cycle? Are there other box scenarios in which the local/noncontextual polytope admits a simple facet characterisation? Can we obtain the efficiency requirements for loophole free Bell tests in more general scenarios from an analytical approach? Is the quantum implementation of nonlocal boxes presented in this dissertation practical for real experiments? Can we propose other physical setups in which its requirements for loophole free Bell tests are within current technology?

We hope that the ideas presented in this text contribute to the study of correlations between physical systems and motivate other researchers to answer some of our questions.

Czas – mój największy wróg, mój najlepszy przyjaciel
Czas – nie używa słów, ale zawsze odnajdzie

Na Krawędzi – Closterkeller

Bibliography

[1] M. T. Quintino, M. Araújo, D. Cavalcanti, M. França Santos, and M. Terra Cunha, “Maximal violations and efficiency requirements for Bell tests with photodetection and homodyne measurements,” *Journal of Physics A Mathematical General* **45**, 21 (2012) 215308, arXiv:1106.2486 [quant-ph].

[2] M. Araújo, M. T. Quintino, D. Cavalcanti, M. F. Santos, A. Cabello, and M. T. Cunha, “Tests of Bell inequality with arbitrarily low photodetection efficiency and homodyne measurements,” *Phys. Rev. A* **86**, 3 (2012) 030101, arXiv:1112.1719 [quant-ph].

[3] C. Teo, M. Araújo, M. T. Quintino, J. Minář, D. Cavalcanti, V. Scarani, M. Terra Cunha, and M. França Santos, “Realistic loophole-free Bell test with atom-photon entanglement,” *Nature Communications* **4**, (2013) , arXiv:1206.0074 [quant-ph].

[4] M. Araújo, M. T. Quintino, C. Budroni, M. T. Cunha, and A. Cabello, “All noncontextuality inequalities for the n-cycle scenario,” *Phys. Rev. A* **88**, 2 (2013) 022118, arXiv:1206.3212 [quant-ph].

[5] W. Hermanns and A. Einstein, *Einstein and the Poet: In Search of the Cosmic Man*. Branden Press, 1983. http://books.google.com.br/books?id=QXCyjj6T5ZUC. Page 58.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004. http://www.stanford.edu/~boyd/cvxbook/.

[7] http://www.tcs.fudan.edu.cn/rudolf/Courses/Algorithms/Alg_ss_07w/Webprojects/Chen_hull/applications.htm.

[8] http://en.wikipedia.org/wiki/Convex.

[9] G. Ziegler, *Lectures on polytopes*. Graduate texts in mathematics. Springer-Verlag, 1995. http://books.google.com.br/books?id=xd25TXSSUcgC.

[10] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, 2006. http://books.google.com.br/books?id=aA4vXMbuOTUC.

[11] N. Lauritzen, "Lectures on Convex Sets." Aarhus university. http://home.imf.au.dk/niels/lecconset.pdf.

[12] "Carathéodory's theorem (convex hull)." Wikipedia. http://en.wikipedia.org/wiki/Carath%C3%A9odory%27s_theorem_%28convex_hull%29.

[13] B. S. Tsirelson, "Some results and problems on quantum Bell-type inequalities," *Hadronic J. Suppl.* **8**, 4 (1993) 329–345. http://www.tau.ac.il/~tsirel/download/hadron.html.

[14] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, 2006. http://books.google.com.br/books?id=EuhBluW31hsC.

[15] J. S. Bell, "On the Einstein-Poldolsky-Rosen paradox," *Physics* **1**, 3 (1964) 195–200.

[16] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000. http://books.google.com.br/books?id=65FqEKQOfP8C.

[17] G. Murta, "Não-localidade em sistemas quânticos," 2012. http://www13.fisica.ufmg.br/~posgrad/Dissertacoes_Mestrado/decada2010/glaucia-murta/GlauciaMGuimaraes-diss.pdf. Master's thesis.

[18] M. T. C. Quintino, "Não-localidade como recurso para a comunicação," 2010. http://www.mat.ufmg.br/~tcunha/MonografiaMTulio.pdf.

[19] S. Pironio, "Aspects of quantum nonlocality," 2004. quic.ulb.ac.be/_media/publications/2004-Thesis-Stefano.pdf.

[20] R. Rabelo, "Não-localidade, matemática e fundamentos," 2010. http://www.mat.ufmg.br/~tcunha/DisRafael.pdf. Master's thesis.

[21] K. F. Pál and T. Vértesi, "Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems," *Phys. Rev. A* **82**, 2 (2010) 022116, arXiv:1006.3032 [quant-ph].

[22] I. Pitowsky, *Quantum probability–quantum logic*. Lecture notes in physics. Springer-Verlag, 1989. http://books.google.com.br/books?id=1I4fAQAAMAAJ.

[23] R. F. Werner and M. M. Wolf, "Bell inequalities and entanglement," *Quant. Inf. and Comp.* **3**, (2001) 1, quant-ph/0107093.

[24] K. F. Pál and T. Vértesi, "Concavity of the set of quantum probabilities for any given dimension," *Phys. Rev. A.* **80**, 4 (2009) 042114, arXiv:0905.2915 [quant-ph].

[25] G. Svetlichny, "Distinguishing three-body from two-body nonseparability by a Bell-type inequality," *Phys. Rev. D* **35**, (1987) 3066–3069.

[26] M. Seevinck and G. Svetlichny, "Bell-Type Inequalities for Partial Separability in N-Particle Systems and Quantum Mechanical Violations," *Phys. Rev. Lett.* **89**, (2002) 060401.

[27] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, "Bell-Type Inequalities to Detect True n-Body Nonseparability," *Phys. Rev. Lett.* **88**, (2002) 170405.

[28] J. Barrett, S. Pironio, J.-D. Bancal, and N. Gisin, "The definition of multipartite nonlocality," `arXiv:1112.2626 [quant-ph]`.

[29] G. Boole, "On the Theory of Probabilities," *Philosophical Transactions of the Royal Society of London* **152**, (1862) pp. 225–252. `http://www.jstor.org/stable/108830`.

[30] J. B. Fourier, "Solution d'une Question Particulière du Calcul des Inégalités," *Nouveau Bulletin des Sciences par la Socieété Philomathique de Paris* (1826) 99.

[31] T. Motzkin, "Beitrage zur Theorie der linearen Ungleichungen.,". `http://worldcat.org/oclc/79382233`.

[32] `http://www.ifor.math.ethz.ch/~fukuda/cdd_home/index.html`.

[33] `http://www.iwr.uni-heidelberg.de/groups/comopt/software/PORTA`.

[34] `http://www.ifor.math.ethz.ch/~fukuda/polyfaq/node18.html`.

[35] `http://www.ifor.math.ethz.ch/~fukuda/polyfaq/polyfaq.html`.

[36] `http://en.wikipedia.org/wiki/Computational_complexity_theory`.

[37] M. Froissart, "Constructive generalization of Bell's inequalities," *Il Nuovo Cimento B (1971-1996)* **64**, (1981) 241–251.

[38] G. Boole, *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities*. 1854. `http://books.google.com.br/books?id=DqwAAAAAcAAJ`.

[39] I. Pitowsky, "George Boole's "Conditions of Possible Experience" and the Quantum Puzzle," *The British Journal for the Philosophy of Science* **45**, 1 (1994) 95–125.

[40] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories," *Phys. Rev. Lett.* **23**, 15 (1969) 880–884.

[41] J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*. Addison-Wesley, 2010. http://books.google.ch/books?id=N4I-AQAACAAJ.

[42] A. Fine, "Hidden Variables, Joint Probability, and the Bell Inequalities," *Phys. Rev. Lett.* **48**, (1982) 291–295.

[43] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, "Nonlocal correlations as an information-theoretic resource," *Phys. Rev. A.* **71**, 2 (2005) 022101, arXiv:quant-ph/0404097.

[44] D. Avis and K. Fukuda, "A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra," *Discrete and Computational Geometry* **8**, (1992) 295–313.

[45] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," *Foundations of Physics* **24**, (1994) 379–385.

[46] P. Rastall, "Locality, Bell's theorem, and quantum mechanics," *Foundations of Physics* **15**, (1985) 963–972.

[47] L. A. Khalfin and B. S. Tsirelson, "Quantum and quasi-classical analogs of Bell inequalities," *Symposium on the Foundations of Modern Physics* (1985) 441–460. http://www.tau.ac.il/~tsirel/download/khts85.html.

[48] B. S. Cirel'son, "Quantum generalizations of Bell's inequality," *Letters in Mathematical Physics* **4**, (1980) 93–100.

[49] M. Navascués, S. Pironio, and A. Acín, "Bounding the Set of Quantum Correlations," *Phys. Rev. Lett.* **98**, 1 (2007) 010401, arXiv:quant-ph/0607119.

[50] M. Navascués, S. Pironio, and A. Acín, "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations," *New Journal of Physics* **10**, 7 (2008) 073013, arXiv:0803.4290 [quant-ph].

[51] R. F. Werner and M. M. Wolf, "All-multipartite Bell-correlation inequalities for two dichotomic observables per site," *Phys. Rev. A.* **64**, 3 (2001) 032112, arXiv:quant-ph/0102024.

[52] M. Żukowski and Č. Brukner, "Bell's Theorem for General N-Qubit States," *Phys. Rev. Lett.* **88**, 21 (2002) 210401, arXiv:quant-ph/0102039.

[53] I. Pitowsky, "Geometry of quantum correlations," *Phys. Rev. A.* **77**, 6 (2008) 062109, arXiv:0802.3632 [quant-ph].

[54] L. Landau, "Empirical two-point correlation functions," *Foundations of Physics* **18**, (1988) 449–460.

[55] L. Masanes, "Necessary and sufficient condition for quantum-generated correlations," arXiv:quant-ph/0309137.

[56] S. Abramsky and A. Brandenburger, "The sheaf-theoretic structure of non-locality and contextuality," *New Journal of Physics* **13**, 11 (2011) 113036, arXiv:1102.0264 [quant-ph].

[57] T. Fritz, "Polyhedral duality in Bell scenarios with two binary observables," arXiv:1202.0141 [quant-ph].

[58] E. Schrödinger, "Discussion of Probability Relations between Separated Systems," *Proceedings of the Cambridge Philosophical Society* **31**, (1935) 555.

[59] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Phys. Rev. A* **40**, (1989) 4277–4281.

[60] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.* **69**, (1992) 2881–2884.

[61] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.* **70**, (1993) 1895–1899.

[62] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *eprint arXiv:quant-ph/9508027* (1995) , arXiv:quant-ph/9508027.

[63] N. Gisin, "Bell's inequality holds for all non-product states," *Physics Letters A* **154**, (1991) 201–202.

[64] J. Barrett, "Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality," *Phys. Rev. A.* **65**, 4 (2002) 042302, arXiv:quant-ph/0107045.

[65] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, "Noise Robustness of the Nonlocality of Entangled Quantum States," *Phys. Rev. Lett.* **99**, 4 (2007) 040403, arXiv:quant-ph/0703018.

[66] T. Heinosaari and M. M. Wolf, "Nondisturbing quantum measurements," *Journal of Mathematical Physics* **51**, 9 (2010) 092201, arXiv:1005.5659 [quant-ph].

[67] L. Masanes, "Tight Bell inequality for d-outcome measurements correlations," arXiv:quant-ph/0210073.

[68] J. F. Clauser and M. A. Horne, "Experimental consequences of objective local theories," *Phys. Rev. D* **10**, 2 (1974) 526–535.

[69] J. Watrous, "Theory of Quantum Information." http://www.cs.uwaterloo.ca/~watrous/quant-info/lecture-notes/02.pdf. University of Waterloo, 2008.

[70] L. J. Landau, "On the violation of Bell's inequality in quantum theory," *Phys. Lett. A* **120**, 2 (1987) 54 – 56.

[71] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, "Measurements Incompatible in Quantum Theory Cannot Be Measured Jointly in Any Other No-Signaling Theory," *Phys. Rev. Lett.* **103**, 23 (2009) 230402, arXiv:0905.2998 [quant-ph].

[72] D. Dieks, "Inequalities that test locality in quantum mechanics," *Phys. Rev. A* **66**, 6 (2002) 062104, quant-ph/0206172.

[73] I. Pitowsky and K. Svozil, "Optimal tests of quantum nonlocality," *Phys. Rev. A* **64**, (2001) 014102.

[74] C. Śliwa, "Symmetries of the Bell correlation inequalities," *Physics Letters A* **317**, (2003) 165–168, arXiv:quant-ph/0305190.

[75] D. Collins and N. Gisin, "A relevant two qubit Bell inequality inequivalent to the CHSH inequality," *Journal of Physics A Mathematical General* **37**, (2004) 1775–1787, arXiv:quant-ph/0306129.

[76] A. Peres, "All the Bell inequalities," arXiv:quant-ph/9807017.

[77] A. Peres, "Separability Criterion for Density Matrices," *Phys. Rev. Lett.* **77**, (1996) 1413–1415, arxiv:quant-ph/9604005.

[78] W. Dür, "Multipartite Bound Entangled States that Violate Bell's Inequality," *Phys. Rev. Lett.* **87**, 23 (2001) 230402, arXiv:quant-ph/0107050.

[79] T. Vértesi and N. Brunner, "Quantum Nonlocality Does Not Imply Entanglement Distillability," *Phys. Rev. Lett.* **108**, 3 (2012) 030403, arXiv:1106.4850 [quant-ph].

[80] S. Pironio, J.-D. Bancal, and V. Scarani, "Extremal correlations of the tripartite no-signaling polytope," *Journal of Physics A Mathematical General* **44**, 6 (2011) 065303, arXiv:1101.2477 [quant-ph].

[81] M. L. Almeida, J.-D. Bancal, N. Brunner, A. Acín, N. Gisin, and S. Pironio, "Guess Your Neighbor's Input: A Multipartite Nonlocal Game with No Quantum Advantage," *Phys. Rev. Lett.* **104**, 23 (2010) 230404, arXiv:1003.3844 [quant-ph].

[82] A. Acín, M. L. Almeida, R. Augusiak, and N. Brunner, "Guess your neighbour's input: no quantum advantage but an advantage for quantum theory," *ArXiv e-prints* (2012) , arXiv:1205.3076 [quant-ph].

[83] D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going Beyond Bell's Theorem," arXiv:0712.0921 [quant-ph].

[84] N. D. Mermin, "Quantum mysteries revisited," *American Journal of Physics* **58**, 8 (1990) 731–734.

[85] A. Acín, *et al.*, To be published.

[86] A. Cabello, "Quantum nonlocal boxes and nonlocal box swapping," arXiv:1208.1965 [quant-ph].

[87] A. Cabello, "State-independent quantum contextuality and maximum nonlocality," arXiv:1112.5149 [quant-ph].

[88] A. Cabello, "Experimentally Testable State-Independent Quantum Contextuality," *Phys. Rev. Lett.* **101**, 21 (2008) 210401, arXiv:0808.2456 [quant-ph].

[89] G. Brassard, A. Broadbent, and A. Tapp, "Quantum Pseudo-Telepathy," *Foundations of Physics* **35**, (2005) 1877–1907, arXiv:quant-ph/0407221.

[90] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, "Bell Inequalities for Arbitrarily High-Dimensional Systems," *Phys. Rev. Lett.* **88**, 4 (2002) 040404, arXiv:quant-ph/0106024.

[91] S. L. Braunstein and C. M. Caves, "Wringing out better Bell inequalities," *Nucl. Phys. B - Proceedings Supplements* **6**, 0 (1989) 211 – 221.

[92] L. Masanes, A. Acin, and N. Gisin, "General properties of nonsignaling theories," *Phys. Rev. A* **73**, 1 (2006) 012112, arXiv:quant-ph/0508016.

[93] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature* **299**, 5886 (1982) 802–803.

[94] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, "Device-Independent Security of Quantum Cryptography against Collective Attacks," *Phys. Rev. Lett.* **98**, 23 (2007) 230501, arXiv:quant-ph/0702152.

[95] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, "Device-independent quantum key distribution secure against collective attacks," *New Journal of Physics* **11**, 4 (2009) 045021, arXiv:0903.4460 [quant-ph].

[96] S. Pironio, A. Acín, S. Massar, A. B. de La Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, "Random numbers certified by Bell's theorem," *Nature* **464**, (2010) 1021–1024, arXiv:0911.3427 [quant-ph].

[97] C.-E. Bardyn, T. C. H. Liew, S. Massar, M. McKague, and V. Scarani, "Device-independent state estimation based on Bell's inequalities," *Phys. Rev. A* **80**, 6 (2009) 062327, arXiv:0907.2170.

[98] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, "Device-independent certification of entangled measurements," arXiv:1105.3138 [quant-ph].

[99] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Methot, and V. Scarani, "Testing the Hilbert space dimension," arXiv:0802.0760 [quant-ph].

[100] A. Aspect, "Quantum mechanics: To be or not to be local," *Nature* **446**, (2007) 866–867.

[101] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, "Quantum nonlocality in two three-level systems," *Phys. Rev. A* **65**, 5 (2002) 052325, arXiv:quant-ph/0111143.

[102] M. Hendrych, R. Gallego, M. Mičuda, N. Brunner, A. Acín, and J. P. Torres, "Experimental estimation of the dimension of classical and quantum systems," arXiv:1111.1208 [quant-ph].

[103] R. Gallego, N. Brunner, C. Hadley, and A. Acín, "Device-Independent Tests of Classical and Quantum Dimensions," *Phys. Rev. Lett.* **105**, 23 (2010) 230501, arXiv:1010.5064 [quant-ph].

[104] R. Ramanathan, A. Soeda, P. Kurzynski, and D. Kaszlikowski, "Monogamy of contextuality," arXiv:1201.5836 [quant-ph].

[105] A. Klyachko, "Coherent states, entanglement, and geometric invariant theory," quant-ph/0206012.

[106] R. Chaves and T. Fritz, "Entropic approach to local realism and noncontextuality," *Phys. Rev. A* **85**, 3 (2012) 032113.

[107] T. Fritz and R. Chaves, "Entropic inequalities and the marginal problem," arXiv:1112.4788 [cs.IT].

[108] M. Araújo, "Quantum realism and quantum surrealism," 2012. http://arxiv.org/abs/1208.6283.

[109] http://www.tau.ac.il/~tsirel/Research/bellopalg/main.html.

[110] http://web.archive.org/web/20090607205354/http://www.imaph.tu-bs.de/qi/problems/.

[111] O. Krueger and R. F. Werner, "Some Open Problems in Quantum Information Theory," *eprint arXiv:quant-ph/0504166* (2005) , arXiv:quant-ph/0504166.

[112] http://www.tau.ac.il/~tsirel/download/bellopalg.pdf.

[113] M. Junge, M. Navascues, C. Palazuelos, D. Perez-Garcia, V. B. Scholz, and R. F. Werner, "Connes' embedding problem and Tsirelson's problem," *Journal of Mathematical Physics* **52**, 1 (2011) 012102, arXiv:1008.1142 [math-ph].

[114] M. Navascues, T. Cooney, D. Perez-Garcia, and I. Villanueva, "A physical approach to Tsirelson's problem," arXiv:1105.3373 [quant-ph].

[115] G. Dall'Aglio, S. Kotz, G. Salinetti, and p. e. s. a. Università degli studi di Roma "La Sapienza." Dipartimento di statistica, *Advances in Probability Distributions With Given Marginals: Beyond the Copulas*. Mathematics and Its Applications. Kluwer Academic Publishers, 1991. http://books.google.com.br/books?id=4uNCdVrrw2cC.

[116] L. Rüschendorf, B. Schweizer, and M. Taylor, *Distributions with Fixed Marginals and Related Topics*. Lecture notes-monograph series. Institute of Mathematical Statistics, 1996.
http://books.google.com.br/books?id=4FkPd0IgKkwC.

[117] V. Beneš and J. Štěpán, *Distributions with given marginals and moment problems*. Kluwer Academic Publishers, 1997.
http://books.google.com.br/books?id=YQzvAAAAMAAJ.

[118] C. Budroni and G. Morchio, "The extension problem for partial Boolean structures in quantum mechanics," *Journal of Mathematical Physics* **51**, 12 (2010) 122205, arXiv:1010.4662 [quant-ph].

[119] E. P. Specker, "Die Logik nicht gleichzeitig entscheidbarer Aussagen," *Dialectica* **14**, 2-3 (1960) 239–246.

[120] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, "Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity," *Physics Reports* (2011) 1–39.

[121] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, "Simple Test for Hidden Variables in Spin-1 Systems," *Phys. Rev. Lett.* **101**, 2 (2008) 020403, arXiv:0706.0126 [quant-ph].

[122] A. Cabello, S. Severini, and A. Winter, "(Non-)Contextuality of physical theories as an axiom," arXiv:1010.2163 [quant-ph].

[123] S. Wehner, "Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities," *Phys. Rev. A* **73**, 2 (2006) 022110.

[124] I. Pitowsky, "Most Bell Operators do not Significantly Violate Locality," arXiv:quant-ph/0202053.

[125] R. Drumond and R. Oliveira, "Small violations of full correlation Bell inequalities for multipartite pure random states.". To be published.

[126] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum mechanics*. Quantum Mechanics. Wiley, 1977.
http://books.google.com.br/books?id=CnkfAQAAMAAJ.

[127] M. T. Cunha, *Noções de Informação Quântica*. IMPA, 2007.
http://www.mat.ufmg.br/~tcunha/InfoQuanMTerraCunha.pdf.

[128] M. T. Cunha, B. Amaral, and A. Baraviera, *Mecânica Quântica para Matemáticos em Formação*. IMPA, 2011.
http://www.mat.ufmg.br/~tcunha/28CBM_MQM.pdf.

[129] http://en.wikipedia.org/wiki/Spectral_theorem.

[130] http://en.wikipedia.org/wiki/Pauli_matrices.

[131] S. J. Freedman and J. F. Clauser, "Experimental Test of Local Hidden-Variable Theories," *Phys. Rev. Lett.* **28**, (1972) 938–941.

[132] A. Aspect, J. Dalibard, and G. Roger, "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers," *Phys. Rev. Lett.* **49**, 25 (1982) 1804–1807.

[133] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, "New High-Intensity Source of Polarization-Entangled Photon Pairs," *Phys. Rev. Lett.* **75**, (1995) 4337–4341.

[134] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of Bell's Inequality under Strict Einstein Locality Conditions," *Phys. Rev. Lett.* **81**, (1998) 5039–5043.

[135] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, "Experimental violation of a Bell's inequality with efficient detection," *Nature* **409**, (2001) 791.

[136] D. N. Matsukevich, P. Maunz, D. L. Moehring, S. Olmschenk, and C. Monroe, "Bell Inequality Violation with Two Remote Atomic Qubits," *Phys. Rev. Lett.* **100**, (2008) 150404.

[137] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.-S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N. K. Langford, T. Jennewein, and A. Zeilinger, "Violation of local realism with freedom of choice," *PNAS* **107**, 46 (2010) 19708–19713.

[138] Z. Merali, "Quantum Mechanics Braces for the Ultimate Test," *Science* **331**, 6023 (2011) 1380–1382.

[139] C. Branciard, "Detection Loophole in Bell experiments: How post-selected local correlations can look non-local," `arXiv:1010.1178 [quant-ph]`.

[140] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995. `http://books.google.com.br/books?id=FeBix14iM70C`.

[141] G. Garbarino, "Minimum detection efficiencies for a loophole-free observable-asymmetric Bell-type test," *Phys. Rev. A* **81**, 3 (2010) 032106, `arXiv:0912.2258 [quant-ph]`.

[142] J.-A. Larsson and J. Semitecolos, "Strict detector-efficiency bounds for n-site Clauser-Horne inequalities," *Phys. Rev. A* **63**, 2 (2001) 022117, `quant-ph/0006022`.

[143] A. Cabello and J.-Å. Larsson, "Minimum Detection Efficiency for a Loophole-Free Atom-Photon Bell Experiment," *Phys. Rev. Lett.* **98**, 22 (2007) 220402, `arXiv:quant-ph/0701191`.

[144] P. H. Eberhard, "Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment," *Physical Rev. A* **47**, 2 (1993) 747–750.

[145] D. Cavalcanti, N. Brunner, P. Skrzypczyk, A. Salles, and V. Scarani, "Large violation of Bell inequalities using both particle and wave measurements," *Phys. Rev. A* **84**, (2011) 022105, arXiv:1012.1916 [quant-ph].

[146] S.-W. Ji, J. Kim, H.-W. Lee, M. S. Zubairy, and H. Nha, "Loophole-Free Bell Test for Continuous Variables via Wave and Particle Correlations," *Phys. Rev. Lett.* **105**, 17 (2010) 170404, arXiv:1010.4587 [quant-ph].

[147] D. Cavalcanti and V. Scarani, "Comment on "Loophole-Free Bell Test for Continuous Variables via Wave and Particle Correlations"," *Phys. Rev. Lett.* **106**, 20 (2011) 208901, arXiv:1010.5358 [quant-ph].

[148] A. Gilchrist, P. Deuar, and M. D. Reid, "Contradiction of Quantum Mechanics with Local Hidden Variables for Quadrature Phase Amplitude Measurements," *Phys. Rev. Lett.* **80**, 15 (1998) 3169–3172.

[149] W. J. Munro, "Optimal states for Bell-inequality violations using quadrature-phase homodyne measurements," *Phys. Rev. A* **59**, 6 (1999) 4197–4201.

[150] J. Wenger, M. Hafezi, F. Grosshans, R. Tualle-Brouri, and P. Grangier, "Maximal violation of Bell inequalities using continuous-variable measurements," *Phys. Rev. A* **67**, 1 (2003) 012105, quant-ph/0211067.

[151] E. G. Cavalcanti, C. J. Foster, M. D. Reid, and P. D. Drummond, "Bell Inequalities for Continuous-Variable Correlations," *Phys. Rev. Lett.* **99**, 21 (2007) 210405, arXiv:0705.1385 [quant-ph].

[152] A. Acín, N. J. Cerf, A. Ferraro, and J. Niset, "Tests of multimode quantum nonlocality with homodyne measurements," *Phys. Rev. A* **79**, 1 (2009) 012112, arXiv:0808.2373 [quant-ph].

[153] Q. Y. He, E. G. Cavalcanti, M. D. Reid, and P. D. Drummond, "Bell inequalities for continuous-variable measurements," *Phys. Rev. A* **81**, 6 (2010) 062106, arXiv:1005.2208 [quant-ph].

[154] H. Nha and H. J. Carmichael, "Proposed Test of Quantum Nonlocality for Continuous Variables," *Phys. Rev. Lett.* **93**, 2 (2004) 020401.

[155] R. García-Patrón, J. Fiurášek, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, "Proposal for a Loophole-Free Bell Test Using Homodyne Detection," *Phys. Rev. Lett.* **93**, 13 (2004) 130409.

[156] http://en.wikipedia.org/wiki/Hermite_polynomials.

[157] http://en.wikipedia.org/wiki/Data_binning.

[158] M. Reed and B. Simon, *Methods of Modern Mathematical Physics: Functional analysis*. No. v. 1 in Methods of Modern Mathematical Physics. Academic Press, 1980. http://books.google.com.br/books?id=fXX0j4qa8G8C.

[159] N. Sangouard, J.-D. Bancal, N. Gisin, W. Rosenfeld, P. Sekatski, M. Weber, and H. Weinfurter, "Loophole-free Bell test with one atom and less than one photon on average," *Phys. Rev. A* **84**, 5 (2011) 052122, arXiv:1108.1027 [quant-ph].

[160] H. Jeong and T. C. Ralph, "Schrödinger Cat States for Quantum Information Processing," arXiv:quant-ph/0509137.

[161] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri1, and P. Grangier, "Generation of optical "Schrodinger cats" from photon number states," *Nature* **448**, (2007) 784–786.

[162] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," *Phys. Rev.* **131**, (1963) 2766–2788. http://link.aps.org/doi/10.1103/PhysRev.131.2766.

[163] J. M. Raimond, M. Brune, and S. Haroche, "Manipulating quantum entanglement with atoms and photons in a cavity," *Rev. Mod. Phys.* **73**, (2001) 565–582. http://link.aps.org/doi/10.1103/RevModPhys.73.565.

[164] S. Ritter, C. Nölleke, C. Hahn, A. Reiserer, A. Neuzner, M. Uphoff, M. Mücke, E. Figueroa, J. Bochmann, and G. Rempe, "An elementary quantum network of single atoms in optical cavities," *Nature* **484**, (2012) 195–200, arXiv:1202.5955 [quant-ph].

[165] K. M. Birnbaum, A. Boca, R. Miller, A. D. Boozer, T. E. Northup, and H. J. Kimble, "Photon blockade in an optical cavity with one trapped atom," *Nature* **436**, (2005) 87–90, arXiv:quant-ph/0507065.

[166] T. Vértesi, S. Pironio, and N. Brunner, "Closing the Detection Loophole in Bell Experiments Using Qudits," *Phys. Rev. Lett.* **104**, 6 (2010) 060401, arXiv:0909.3171 [quant-ph].

[167] K. F. Pal, T. Vertesi, and N. Brunner, "Closing the detection loophole in multipartite Bell tests using GHZ states," arXiv:1208.0622 [quant-ph].

[168] R. Chaves and J. B. Brask, "Feasibility of loophole-free nonlocality tests with a single photon," *Phys. Rev. A.* **84**, 6 (2011) 062110, arXiv:1110.3721 [quant-ph].